

\*\*\*\*\*  
 WISE (TM)  
 \*\*\*\*\*

Release 3.1a John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Run on: Wed Aug 16 09:32:39 2000; MasPar time 20.12 Seconds

Tabular output not generated.

Title: >US-09-416-267-2  
 Description: (1-223) from US09416267.pep  
 Perfect Score: 1738  
 Sequence: 1 MKLHVAVLTALILMFLTWL.....IGPCIDYSGKTVKCMCMF 223

Scoring table: PAM 150  
 Gap 11

Searched: 225878 segs, 69334122 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database:

sptrembl12  
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 43.035; Variance 77.012; scale 0.555

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	102	5.9	417	10	664618	6.42e-02
2	103	5.9	434	11	089030	4.57e-02
3	102	5.9	640	11	009182	6.42e-02
4	100	5.8	417	4	09Y409	1.26e-01
5	100	5.8	633	5	022468	1.26e-01
6	98	5.6	1360	10	092VU3	2.45e-01
7	95	5.5	439	4	016042	6.54e-01
8	95	5.5	596	4	095612	6.54e-01
9	94	5.4	246	4	099750	9.02e-01
10	93	5.4	756	5	044417	1.24e+00
11	93	5.4	1454	14	066928	1.24e+00
12	94	5.4	2100	5	021281	9.02e-01
13	92	5.3	1096	3	014022	1.71e+00
14	90	5.2	40	14	057150	3.19e+00
15	90	5.2	116	2	09X2C3	3.19e+00
16	90	5.2	157	11	062331	3.19e+00
17	90	5.2	241	5	022205	3.19e+00
18	90	5.2	343	13	042607	3.19e+00
19	91	5.2	618	3	006817	2.34e+00
20	91	5.2	690	3	006993	2.34e+00

## ALIGNMENTS

RESULT ID	1	PRELIMINARY:	PRT:	417 AA.	SPECIAL LOBE-SPECIFIC PUTATIVE MEMBRANE-ASSO	2.34e+00
AC	064618				MCOP HOMOLOG.	3.19e+00
DT	01-AUG-1998 (TREMBLrel. 07, Created)					
DT	01-NOV-1998 (TREMBLrel. 07, Last sequence update)					
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)					
DE	F19P24.8 PROTEIN.					
GN	F19P24.8.					
OS	Arabidopsis thaliana (Mouse-ear cross).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;					
OC	core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;					
OC	Arabidopsids.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=CV. COLUMBIA;					
RA	ROUSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,					
RA	SYKES S.M., KAUL S., MASON T.M., KERLAJAGE A.R., ADAMS M.D.,					
RA	SOMERVILLE C.R., VENTER J.C.;					
RL	Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.					
DR	EMBL: AC003673; AAC09024.1;					
DR	MEDEL: 28581; Arath:3412;28581.					
DR	PFAM: PF00041; fn3; 1.					
SO	SEQUENCE 417 AA; 47607 MW; 6A31B808 CRC32;					
Query Match	5.9%	Score 102;	DB 10;	Length 417;		
Best Local Similarity	27.5%	Pred. No. 6.42e-02;				
Matches	22;	Conservative 20;	Mismatches 34;	Indels 4;	Gaps 4;	
DB	273 WFKVFSGEVSELSDEFIVS-TKTLDEEVAVALNNSNCNNAKMKSGSCSPG-PEE 330					
QY	116 W-NIVSFPAEELSHENLVSLFETVNPQPHQNVSPSNNAHAPRSSDKHEHCIVYIEDD 174					
DB	331 CYNLIR-QLECSGQVKSDFR 349					
QY	175 CSMHCKSCSCSMGASKYR 194					
RESULT 2						
ID	089030	PRELIMINARY:	PRT:	434 AA.		
AC	089030;					
DT	01-NOV-1998 (TREMBLrel. 08, Created)					
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)					
DT	01-NOV-1999 (TREMBLrel. 12, Last annotation update)					

DE MATRILIN-4 PRECURSOR, ALTERNATE SPLICE PRODUCT.  
 GN MAT-4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J; TISSUE-TOTAL FOETUS;  
 RX MEDLINE: 98442849.  
 RA WAGNER R., KOBBE B., PAULSON M.;  
 RT Matrilin-4, a new member of the matrilin family of extracellular  
 matrix proteins.  
 RL FEBS Lett. 436:123-127(1998).  
 DR EMBL: AJ006140; CAA06890.1; .  
 DR HSSP: P00736; IAP0.  
 DR PROSITE: PS001010; ASX\_HYDROXYL; 2.  
 DR PFAM: PF00008; EGF; 4.  
 DR PFAM: PF00092; vwa; 1.  
 DR PRINTS: PR00453; VWFADOMAIN.  
 KW Signal; Glycoprotein; EGF-like domain.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 434 MATRILIN-4 PRECURSOR.  
 FT CHAIN 22 434 MATRILIN-4 PRECURSOR, ALTERNATE SPLICE  
 FT VARIANT 129 129 PRODUCT.  
 FT VARIANT 156 156 A -> E.  
 FT VARIANT 156 156 G -> D.  
 SQ SEQUENCE 434 AA; 47944 MW; B29D569A CRC32;  
 Query Match 5.9%; Score 103; DB 11; Length 434;  
 Best Local Similarity 41.7%; Pred. No. 4,576-02;  
 Matches 20; Conservative 6; Mismatches 18; Indels 4; Gaps 3;  
 Db 9 LSL-LTFLFOSWETOLOSAGKDLCAELVHG--OHLGNAPGTFYAC 53  
 QY 9 LTLALIMFL-TWLPESLSCNKLACSDVSKLIDELCCCRPGNGSC 55  
 RESULT 3  
 AC 009182; PRELIMINARY; PRT; 640 AA.  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE INTEGRIN BETA-7 SUBUNIT (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-SPRAGUE-DAWLEY; TISSUE-SPLEEN;  
 RT TIDSWELL M., PACHYNSKI R., WU S.W., QIU S.O., DUNHAM E., COCHRAN N.,  
 BRISKIN M.J., KILSHAW P.J., LAZAROVITS A.I., ANDREW D.P.,  
 BUTCHER E.C., YEDNOCK T.A., ERLE D.J.;  
 RL J. Immunol. 0:0-0(0).  
 CC -1- FUNCTION: INTEGRINS ARE A LARGE FAMILY OF CELL SURFACE  
 GLYCOPROTEINS THAT MEDIANTE CELL TO CELL & CELL TO MATRIX ADHESION.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: WITH OTHER BETA CHAINS FROM THE INTEGRIN FAMILY OF  
 CELL-SURFACE RECEPTOR.  
 DR EMBL: AF003598; AAB61241.1; .  
 DR PROSITE: PS00243; INTEGRIN\_BETA; 3.  
 DR PFAM: PF00362; Integrin\_B; 1.  
 KW Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat.  
 FT NON\_TER 1  
 SQ SEQUENCE 640 AA; 69015 MW; 313C3A97 CRC32;  
 Query Match 5.9%; Score 102; DB 11; Length 640;  
 Best Local Similarity 31.6%; Pred. No. 6,426-02;  
 Matches 18; Conservative 13; Mismatches 22; Indels 4; Gaps 4;  
 Db 432 CSEVSDSCVSEGGIC-SGHGDKCNRCQCLDGYGALCDCLCKSPCYRDCAE 487  
 QY 31 CASDVSKLIDELCCCRPGNGSCCK-ECN-LCLGALIMDECCCYGMCP-RKYS 84

RESULT 4  
 ID 09409; PRELIMINARY; PRT; 417 AA.  
 AC 09409;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE HYPOTHETICAL 44.9 KD PROTEIN.  
 GN DKF2P566D213.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homidae; Homo.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-KIDNEY;  
 RA KOEHLER K., BEYER A., MEMES H.W., GASSENHUBER J., WIEBMAN S.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL050275; CAB43376.1; .  
 DR PROSITE: PS01187; EGF\_CA; 2.  
 KW Hypothetical protein; Glycoprotein; EGF-like domain.  
 SQ SEQUENCE 417 AA; 44934 MW; 96417E05 CRC32;  
 Query Match 5.8%; Score 100; DB 4; Length 417;  
 Best Local Similarity 28.6%; Pred. No. 1,266-01;  
 Matches 14; Conservative 12; Mismatches 21; Indels 2; Gaps 2;  
 Db 236 WALHLKCYDIDECGEGANGCAGDQCVNTGSSYECRCACACACGMA 284  
 QY 19 WLPESLSC-NKLACSDVSKLIDELCCCRPGNGSC-SCRECMCLCA 65  
 RESULT 5  
 ID 022468; PRELIMINARY; PRT; 633 AA.  
 AC 022468;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)  
 DE T13H5.3 PROTEIN.  
 GN T13H5.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditidae; Rhabditidae; Pelodidae; Caenorhabditis.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP LIGHTNING J.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE: 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., CORSEY T., COOPER J., COULSON A.,  
 RA CRATON M., DEAR S., DU Z., DUBBIN R., FAVELLO A., FULTON L.,  
 RA GARNER A., GREEN P., HAWKINS T., HILLER L., JIR M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STRADEN R., SUTLSON J.,  
 RA THIRRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLIDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans."  
 RL Nature 368:32-38(1994).  
 DR EMBL: 26524; CAA91419.1; .  
 DR HSSP: P27275; IMMC.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 DR PFAM: PF00704; Glyco\_hydro\_18; 1.  
 KW Hydrolase; Glycosidase.  
 SQ SEQUENCE 633 AA; 70746 MW; F23278E2 CRC32;  
 Query Match 5.8%; Score 100; DB 5; Length 633;  
 Best Local Similarity 26.0%; Pred. No. 1,266-01;  
 Matches 19; Conservative 18; Mismatches 31; Indels 5; Gaps 5;

Db 494 ITWTSADAGEKRCGRCDVPP-LEGEAPTCNPDANAHCCSNGCYC-GNSKEHCENGC 551  
 17 LFWL-PELSKCNKALCASDVSKLIDELCCRCRPGECNCSCKEOMLGLMDECC-CD-C 73

Db 552 IDFAKORDFKYK 564  
 74 VGMCPNRYSDPT 86

RESULT 6 PRELIMINARY: PRT: 1360 AA.

AC 092VU3;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DE T5A14.15 PROTEIN.

OC Arabidopsis thaliana (Mouse-ear cress).  
 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;  
 OC Arabidopsis.

RA [1]  
 RA SEQUENCE FROM N.A.  
 RA FEDERSPIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,  
 RA ALTAFT H., ARAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., DUNN P.,  
 RA GONZALEZ A., KREMENTSKAYA I., KIM C., LENZ C., LI J., LIU S.,  
 RA LUKOS S., SCHWARTZ J., SHINN P., TORIUMI M., VYSOTSKAIA V.S.,  
 RA WALKER M., XU G., ECKER J., THEOLOGIS A., DAVIS R.W.,  
 RA Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AC005223; AADI0650.1;  
 SR SEQUENCE 1360 AA; 154661 MW; A7040DF1 CRC32;

Query Match 5.6%; Score 98; DB 10; Length 1360;  
 Best Local Similarity 31.9%; Pred. No. 2.45e-01;  
 Matches 22; Conservative 12; Mismatches 29; Indels 6; Gaps 6;

Db 1218 FLECIKCKSFGLKATLPSEAHYRHRPL-TLCY-RDYMTSHHWCET-CESK-LSITY 1273  
 136 FLEIYNQPHQNVSPSNVNAAPYSSDKENHCTVYFPDCKSHQ-CIKSESGAKSKYR 194

Db 1274 WFY-TCDSK 1281  
 195 WFNACCEC 203

RESULT 7 PRELIMINARY: PRT: 439 AA.

AC 016042;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE TUMOR NECROSIS FACTOR RECEPTOR.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RA [1]  
 RA SEQUENCE FROM N.A.  
 RA MEDLINE: 91370690.  
 RA DEMBIC Z., LOETSCHER H., GUBLER U., PAN Y.C., LAHM H.W., GENTZ R.,  
 RA BROCKHAUS M., LESSLAUER W.,  
 RA "Two human TNF receptors have similar extracellular, but distinct  
 RA intracellular, domain sequences."  
 RA Cytokine 2:231-237(1990).

DR EMBL: S63368; AAB19824.1;  
 DR HSBP: P25942; ICDP.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; 2.  
 DR PFM: PF00020; TNFR\_C6; 4.  
 SR SEQUENCE 439 AA; 46090 MW; 1CE766FB CRC32;

Query Match 5.5%; Score 95; DB 4; Length 439;  
 Best Local Similarity 42.4%; Pred. No. 6.54e-01;

Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;  
 Db 67 WPEELSCG-SRCSDDOYETQACTREONRICTRPG-WYCALSKQEGCRLC-APLRK-C 121  
 19 WPELSKCNKALCASD-V-SK-CLI-QE-LCCRCRPGECNCSCK-E-CMLCGLMDECC 70

RESULT 8 PRELIMINARY: PRT: 596 AA.

AC P95612;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
 DE MODULATION PROTEIN.  
 GN NOLX.

OS Rhizobium fredii.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.

RA [1]  
 RA SEQUENCE FROM N.A.  
 RA STRAIN-USDA191;  
 RA BELLATO C., KRISHNAN H.B., CUBO T., VERA F.T., PUEPPRE S.G.,  
 RA Microbiology 0:0-0(0).  
 RL EMBL: U77701; AAB19229.1;  
 SR SEQUENCE 596 AA; 64012 MW; CF154622 CRC32;

Query Match 5.5%; Score 95; DB 2; Length 596;  
 Best Local Similarity 29.9%; Pred. No. 6.54e-01;  
 Matches 23; Conservative 21; Mismatches 30; Indels 3; Gaps 3;

Db 156 PSTPPLKAITEFLD-PELFYAGSGDGRGGRKIRAKDSESKHHPQVAAQESQA 214  
 83 SDTPPTSKSTVEELHPIPSLFRAL-TEDGTQLNMNIVSFPVAELSHENLVSLFTVN 141

Db 215 QSYAONY-IPSDSAENA 230  
 142 QPHQNVSPSNVNAAP 158

RESULT 9 PRELIMINARY: PRT: 246 AA.

AC 099750;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE MYOGENIC REPRESSOR I-MF.

GN MDT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RA [1]  
 RA SEQUENCE FROM N.A.  
 RA TISSUE-BRAIN:  
 RA MEDLINE: 96390847.  
 RA CHEN C.M., KRAAT N., GROUNDINE M., WEINTRAUB H.,  
 RA "I-MF, a novel myogenic repressor, interacts with members of the MyoD  
 RA family."  
 RA Cell 86:731-741(1996).  
 RL EMBL: U78313; AAB39748.1;  
 DR EMBL: U78313; AAB39748.1;  
 SR SEQUENCE 246 AA; 25029 MW; 0CD862A7 CRC32;

Query Match 5.4%; Score 94; DB 4; Length 246;  
 Best Local Similarity 38.7%; Pred. No. 9.02e-01;  
 Matches 24; Conservative 8; Mismatches 20; Indels 10; Gaps 9;

Db 175 EPLTICNIVIDCAT-GGSCSSESCICCCCGSGCADDDLCDDC-GIL-DACCESAD 231  
 22 ESLS-CNKAL-CASDVSKLIDF-LCCRCRPGECNCSCK-FCML-CGLMDECCDVG 75

Db 232 -C 232  
 76 MC 77



KW Hypothetical protein; Hydrolyase; Calcium transport; Transmembrane;  
 KM Phosphorylation; Magnesium; ATP-binding.  
 FT TRANSMEM 48 68 POTENTIAL.  
 FT TRANSMEM 209 229 POTENTIAL.  
 FT TRANSMEM 394 414 POTENTIAL.  
 FT TRANSMEM 429 449 POTENTIAL.  
 FT TRANSMEM 893 913 POTENTIAL.  
 FT TRANSMEM 923 943 POTENTIAL.  
 FT TRANSMEM 969 989 POTENTIAL.  
 FT TRANSMEM 1004 1024 POTENTIAL.  
 FT TRANSMEM 1074 1094 POTENTIAL.  
 FT MOD\_RSS 480 480 PHOSPHORYLATION (POTENTIAL).  
 SQ SEQUENCE 1096 AA; 121934 MW; A798E270 CRC32;

Query Match 5.3%; Score 92; DB 3; Length 1096;  
 Best Local Similarity 41.2%; Pred. No. 1.71e+00;  
 Matches 14; Conservative 9; Mismatches 9; Indels 2; Gaps 2;

OY 3 LRVTEILTLGLHLHLTLWLPFRKWIAPCSNE 84  
 1 LHYVAVLTALIL-MELTWLPE-SLSCNACALCSD 34

RESULT 14  
 ID 057150 PRELIMINARY; PRT; 40 AA.  
 AC 057150;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE SIMILAR TO HHV6A U86.  
 GN H88.  
 OS Human herpesvirus 6.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 CC Betaherpesvirinae; Roseolovirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HST;  
 RX MEDLINE; 98115230.  
 RA KOSUGE H., ISEGAWA Y., YAMANISHI K.;  
 RT "Nucleotide sequence analysis of a 30-kilobase-pair region of human  
 RT herpesvirus-6B (HHV-6B) genome and strain-specific variations in major  
 RL immediate-early genes.";  
 DR EMBL; U92288; AAC40337.1; .  
 SEQUENCE 40 AA; 4498 MW; 09D964E5 CRC32;

Query Match 5.2%; Score 90; DB 14; Length 40;  
 Best Local Similarity 50.0%; Pred. No. 3.19e+00;  
 Matches 11; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

DB 12 CVCC-VCVCLCVVF-VCCVCV 31  
 OY 53 CSCCKEMLCIGALMDECCDCV 74

RESULT 15  
 ID 09X2C3 PRELIMINARY; PRT; 116 AA.  
 AC 09X2C3;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE DNAJ-RELATED PROTEIN.  
 GN TM1803.  
 OS Thermotoga maritima.  
 CC Bacteria; Thermotogales; Thermotoga.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 99287316.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,

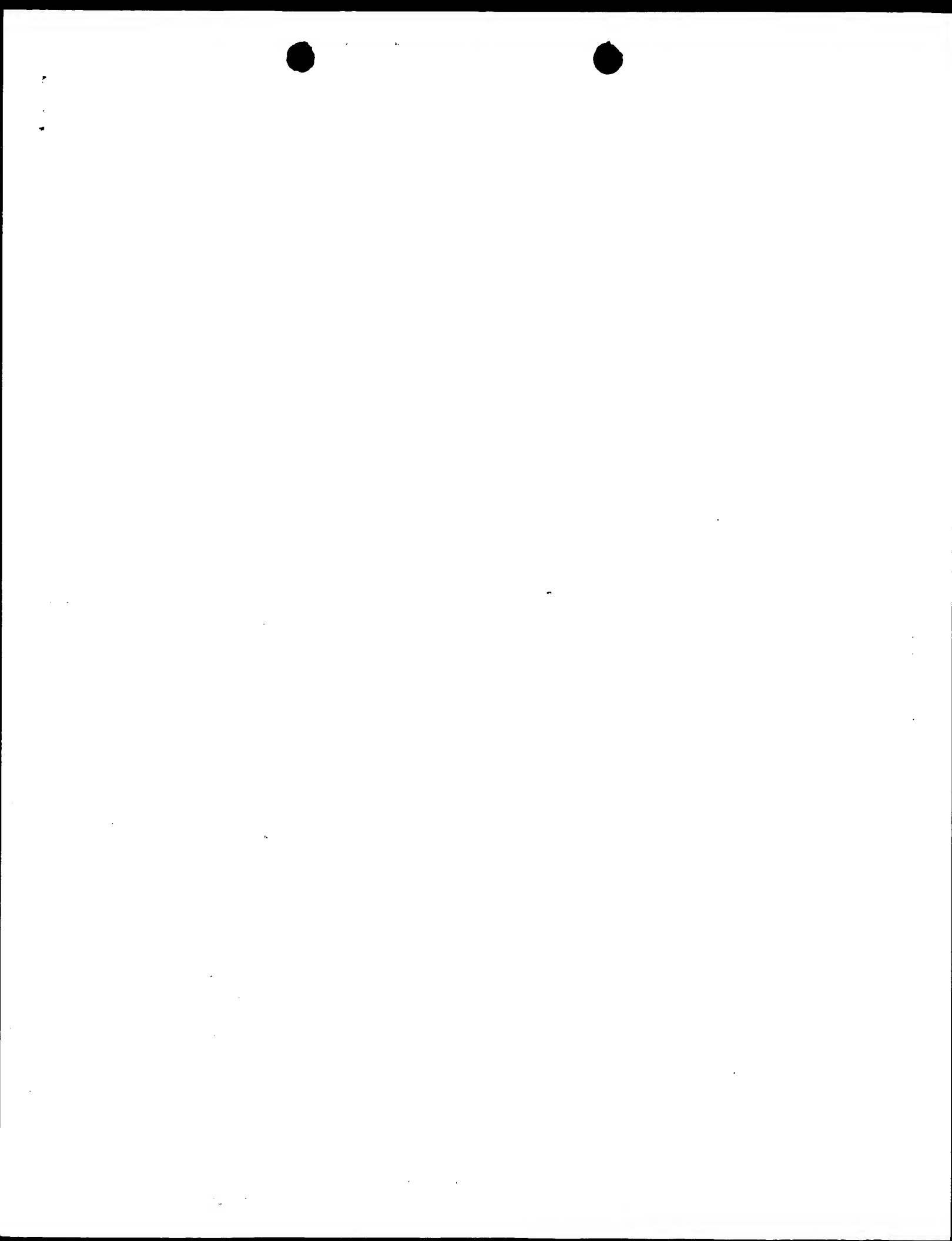
RA SMITH H.O., VENTER J.C., FRASER C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and bacteria from  
 RT genome sequence of Thermotoga maritima.";  
 RL Nature 399:323-329(1999).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GINN M.L., DODSON R.J.,  
 RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,  
 RA MCDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,  
 RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,  
 RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,  
 RA SMITH H.O., VENTER J.C., FRASER C.M.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE001818; AAD36866.1; .  
 SQ SEQUENCE 116 AA; 13100 MW; 4E863268 CRC32;

Query Match 5.2%; Score 90; DB 2; Length 116;  
 Best Local Similarity 44.4%; Pred. No. 3.19e+00;  
 Matches 12; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

DB 88 CRDILACGACACTDCCACGAGCIP 114  
 OY 56 CRCMCLCIGALM--DECCD-CVGMKCP 79

Search completed: Wed Aug 16 09:34:37 2000  
 Job time : 118 secs.



\*\*\*\*\*  
 W P S E R F I  
 (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (C) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

ch\_pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed Aug 16 09:31:40 2000; Maspar time 8.67 Seconds  
 796.748 Million cell updates/sec  
 Tabular output not generated.

Title: >US-09-416-267-2  
 Description: (1-223) from US09416267.pep  
 Perfect Score: 1738  
 Sequence: 1 MKLHVAVLTALIMELTWL.....IGSPECIDYGSVKVCKMCMF 223

Scoring table:  
 PAM 150  
 Gap 11

Searched: 85661 seqs, 3098916 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot38  
 1:swissprot

Statistics: Mean 44.527; Variance 72.235; scale 0.616

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	537	30.9	249	1	TSG_DROME TWISTED GASTRULATION P	4.07e-100
2	104	6.0	798	1	ITB7_HUMAN INTEGRIN BETA-7 PRECUR	7.22e-03
3	101	5.8	194	1	COX2_PARDE PROBABLE CYTOCHROME C	2.16e-02
4	97	5.6	243	1	GTI2_MOUSE GLUTATHIONE S-TRANSFER	8.98e-02
5	97	5.6	806	1	ITB7_MOUSE INTEGRIN BETA-7 PRECUR	8.98e-02
6	97	5.6	1799	1	LMB2_MOUSE LAMININ BETA-2 CHAIN P	8.98e-02
7	96	5.5	326	1	AAIR_BOVIN ADENOSINE A1 RECEPTOR	1.27e-01
8	96	5.5	461	1	AAIR_BOVIN ADENOSINE A1 RECEPTOR	1.27e-01
9	95	5.5	446	1	TNR2_HUMAN TUNOR NECROSIS FACTOR	1.80e-01
10	95	5.5	471	1	NOLX_RHIFR MODULATION PROTEIN NOL	1.80e-01
11	96	5.5	1801	1	LMB2_RAT LAMININ BETA-2 CHAIN P	1.27e-01
12	96	5.5	3084	1	LMB2_MOUSE LAMININ BETA-2 CHAIN P	1.27e-01
13	94	5.4	62	1	MT_XENLA METALLOTHIONEIN	2.54e-01
14	94	5.4	596	1	NOLX_RHIFR MODULATION PROTEIN NOL	2.54e-01
15	94	5.4	3075	1	LMB2_MOUSE LAMININ BETA-2 CHAIN P	2.54e-01
16	92	5.3	130	1	YK06_YEAST HYPOTHETICAL 14.9 KDA	5.02e-01
17	92	5.3	558	1	MEMBRANE TRANSPORTER D	5.02e-01
18	92	5.3	1096	1	ATCY_SCHPO PROBABLE CATION-TRANSP	5.02e-01
19	90	5.2	63	1	MTI_COLLI METALLOTHIONEIN-1 (MT-	9.81e-01
20	90	5.2	167	1	Y199_MYCCE HYPOTHETICAL PROTEIN M	9.81e-01
21	90	5.2	324	1	AAIR_CHICK ADENOSINE A1 RECEPTOR	9.81e-01
22	91	5.2	445	1	CIUS_CANFA CLUSTERIN PRECURSOR (G	7.03e-01
23	90	5.2	690	1	AFI1_YEAST IRON-REGULATED TRANSCR	9.81e-01

RESULT ID	1	STANDARD	PRT	249 AA	ALIGNMENTS
AC	PS4356				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	TWISTED GASTRULATION PROTEIN PRECURSOR				
GN	TSG				
OS	Drosophila melanogaster (Fruit fly)				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-EMBRYO				
RX	MEDLINE: 95047309				
RA	Mason E.D., Konrad K.D., Webb C.D., Marsh J.L.				
RT	"Dorsal midline fate in Drosophila embryos requires twisted gastrulation, a gene encoding a secreted protein related to human connective tissue growth factor."				
RL	Genes Dev. 8:1489-1501(1994)				
CC	-1- FUNCTION: SPECIFY THE FATE OF DORSAL CELLS IN DROSOPHILA EMBRYOS.				
CC	MUTATIONS OF TSG ONLY AFFECT THE FATE OF A NARROW STRIP OF DORSAL				
CC	MIDLINE CELLS AND DO NOT AFFECT DORSAL ECTODERM CELLS.				
CC	-1- SIMILARITY: SOME, TO THE CEF-10/CYR61/CTFG/ETSP-12/NOV PROTEIN				
CC	FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> )				
CC	EMBL: U09808; AAC24234.1; -				
DR	FLYBASE: FBgn003865; tsg.				
DR	Developmental protein; signal.				
KW	STGNL				
FT	CHAIN				
FT	CARBOHYD				
FT	SEQUENCE				
FT	249 AA; 27220 MW; 313566406E4B86D CRC64;				
Query Match	30.98; Score 537; DB 1; Length 249;				
Best Local Similarity	35.48; Pred. No. 4.07e-100;				
Matches	79; Conservative 54; Mismatches 71; Indels 19; Gaps 13;				

DB 1 MOLCYFVILFVGIAPW-SSLANDGNEVSGVSKLITOSCCKRLND--CHCRKDC 57  
 QY 1 MKLH-YVAVLTLLMFLWLPESLSCNALCASDVSKCLIDELCCCRGEGNSCCKEC 59  
 DB 58 LNCGLYELTCCGCLDMC-PKHANDVLPSTLPRSEIGDI-EGVPELFDLTADDE-GMST 114  
 QY 60 MCLGALMDECCDCVGMCPNPNRYSDTRP-TSKSTVELEHPIPSLFALTEBDTQLNMNI 118  
 DB 115 IIRSMAGFOR---V---QG-GAS--GDAGNGNGNAG-SAGVT-LCTVIYVNSCIRA 163  
 QY 119 VSEPVAEELSHENLVSFLETYNQPHONVYVSNVHAPYSSDXEHMCTVYFPDCKSI 178  
 DB 164 NKCRQCESMGASSYRFHDGCEGCEGNCUNYNGINESCRGC 206  
 QY 179 HCKICSESMGASKYRPFHNNACCECIGPECIDIGSKTYACMNC 221  
 RESULT 2  
 ID ITB7\_HUMAN STANDARD: PRT; 798 AA.  
 AC P26010;  
 01-MAY-1992 (Rel. 22, Created)  
 01-MAY-1992 (Rel. 22, Last sequence update)  
 15-FEB-2000 (Rel. 39, Last annotation update)  
 DI INTEGRIN BETA-7 PRECURSOR.  
 GN ITGB7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LEUKOCYTE;  
 RX MEDLINE: 91250405.  
 RA Erie D.J., Ruessig C., Sheppard D., Pytela R.;  
 RT "Complete amino acid sequence of an integrin beta subunit (beta 7)  
 identified in leukocytes."  
 RL J. Biol. Chem. 266:1109-1106(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 91190778.  
 RA Yuan Q., Jiang W.-M., Kissansen G.W., Watson J.D.;  
 RT "Cloning and sequence analysis of a novel beta 2-related integrin  
 transcript from T lymphocytes: homology of integrin cysteine-rich  
 repeats to domain III of laminin B chains."  
 RL Int. Immunol. 2:1097-1108(1990).  
 RN [3]  
 RP REVISIONS.  
 RA Yuan Q., Jiang W.-M., Kissansen G.W., Watson J.D.;  
 RX MEDLINE: 92135083.  
 RA "Cloning and sequence analysis of a novel beta 2-related integrin  
 transcript from T lymphocytes: homology of integrin cysteine-rich  
 repeats to domain III of laminin B chains."  
 RL Int. Immunol. 3:1373-1374(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93002753.  
 RA Jiang W.-M., Jenkins D., Yuan Q., Leung E., Choo K.H., Watson J.D.,  
 RA Kissansen G.W.;  
 RT "The gene organization of the human beta 7 subunit, the common beta  
 subunit of the leukocyte integrins HML-1 and LPAM-1."  
 RL Int. Immunol. 4:1031-1040(1992).  
 CC -1- FUNCTION: EXPECTED TO PLAY A ROLE IN ADHESIVE INTERACTIONS OF  
 CC LEUKOCYTES. INTERACTS WITH ALPHA-4 (TO FORM LPAM-1) OR WITH  
 CC ALPHA-E (TO FORM HML-1).  
 CC -1- SUBUNIT: DIMER OF AN ALPHA AND BETA SUBUNIT. BETA-7 ASSOCIATES  
 CC WITH ALPHA-4 AND -IEL.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- ALTERNATIVE PRODUCTS: THE DIFFERENT FORMS OF BETA-7 SUBUNIT MAY  
 CC ARISE BY ALTERNATIVE SPLICING OF PRIMARY MNA TRANSCRIPTS.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF LEUKOCYTE LINES.  
 CC -1- PAM: THE CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN DISULFIDE  
 CC BONDS.  
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL; S80335; AAB21332.1; -;  
 DR EMBL; M62880; AAA59185.1; -;  
 DR EMBL; M68892; AAA59184.1; -;  
 DR EMBL; S49378; AAB23688.1; -;  
 DR EMBL; S49364; AAB23688.1; JOINED.  
 DR EMBL; S49365; AAB23688.1; JOINED.  
 DR EMBL; S49366; AAB23688.1; JOINED.  
 DR EMBL; S49367; AAB23688.1; JOINED.  
 DR EMBL; S49368; AAB23688.1; JOINED.  
 DR EMBL; S49369; AAB23688.1; JOINED.  
 DR EMBL; S49370; AAB23688.1; JOINED.  
 DR EMBL; S49371; AAB23688.1; JOINED.  
 DR EMBL; S49373; AAB23688.1; JOINED.  
 DR EMBL; S49374; AAB23688.1; JOINED.  
 DR EMBL; S49375; AAB23688.1; JOINED.  
 DR EMBL; S49377; AAB23688.1; JOINED.  
 DR EMBL; L23823; AAA36118.1; -;  
 DR EMBL; L23810; AAA36118.1; JOINED.  
 DR EMBL; L23811; AAA36118.1; JOINED.  
 DR EMBL; L23812; AAA36118.1; JOINED.  
 DR EMBL; L23813; AAA36118.1; JOINED.  
 DR EMBL; L23814; AAA36118.1; JOINED.  
 DR EMBL; L23815; AAA36118.1; JOINED.  
 DR EMBL; L23816; AAA36118.1; JOINED.  
 DR EMBL; L23817; AAA36118.1; JOINED.  
 DR EMBL; L23818; AAA36118.1; JOINED.  
 DR EMBL; L23819; AAA36118.1; JOINED.  
 DR EMBL; L23820; AAA36118.1; JOINED.  
 DR EMBL; L23822; AAA36118.1; JOINED.  
 DR PIR; A40526; A40526.  
 DR MIT; 147559; -;  
 DR PFM; PFM00362; Integrin\_B; 1.  
 DR PRINTS; PRO1186; INTEGRIN.  
 DR PROSITE; PS00243; INTEGRIN\_BETA; 3.  
 DR PROSITE; PS00242; EGF\_1; UNKNOWN\_4.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
 KW Integrin; Cell adhesion; Transmembrane; Glycoprotein; Repeat;  
 KW Extracellular matrix; Cytoskeleton; Signal; Phosphorylation;  
 KW Alternative splicing.  
 FT SIGNAL 1 19  
 FT CHAIN 20 798  
 FT DOMAIN 20 723  
 FT TRANSMEM 724 747  
 FT DOMAIN 747 798  
 FT DOMAIN 747 798  
 FT REPEAT 478 526  
 FT REPEAT 527 565  
 FT REPEAT 566 604  
 FT REPEAT 605 640  
 FT CARBOHYD 68 68  
 FT CARBOHYD 279 279  
 FT CARBOHYD 434 434  
 FT CARBOHYD 477 477  
 FT CARBOHYD 531 531  
 FT CARBOHYD 590 590  
 FT CARBOHYD 655 655  
 FT MOD\_RES 674 674  
 FT VARSPPLIC 501 648  
 FT SEQUENCE 798 AA; 86903 MW; CBE275E0B992385 CMC64;  
 Query Match 6.0%; Score 104; DB 1; Length 798;  
 Best Local Similarity 31.6%; Pred. No. 7, 22e-03;







CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG AND THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ), LAMININ-4  
 CC (S-MEROSIN), AND LAMININ-7 (KS-LAMININ).  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT  
 CC MEMBRANES (MAJOR COMPONENT).  
 CC -1- TISSUE SPECIFICITY: NEUROMUSCULAR SYNAPSE AND KIDNEY GLOMERULUS.  
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.  
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U43541; AAC53535.1; -  
 CC EMBL: U42624; AAC53535.1; JOINED.  
 CC EMBL: X75928; CA53532.1; -  
 CC HSSP: P02468; IKLO.  
 CC MGD: MGI:99916; LAMB2.  
 CC DR PFAM: PF00053; laminin\_EGF\_13.  
 CC DR PRINTS: PRO0011; EGF\_LAMININ.  
 CC DR PROSITE: PS00022; EGF\_2; 1.  
 CC DR PROSITE: PS01186; EGF\_2; 2.  
 CC DR PROSITE: PS01248; LAMININ\_TYPE\_EGF\_12.  
 CC DR Glycoprotein: Basement membrane: Extracellular matrix; Coiled coil;  
 CC Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 CC K1 Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 CC FT SIGNAL 1 35  
 CC FT CHAIN 36 1799 LAMININ BETA-2 CHAIN.  
 CC FT DOMAIN 36 283 LAMININ N-TERMINAL (DOMAIN VI).  
 CC FT DOMAIN 284 555 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN  
 CC V).  
 CC FT 286 349 LAMININ EGF-LIKE 1.  
 CC FT 350 412 LAMININ EGF-LIKE 2.  
 CC FT 413 472 LAMININ EGF-LIKE 3.  
 CC FT 473 524 LAMININ EGF-LIKE 4.  
 CC FT 525 555 LAMININ EGF-LIKE 5 (INCOMPLETE).  
 CC FT 556 782 LAMININ DOMAIN IV.  
 CC FT 783 1191 8 X LAMININ EGF-LIKE REPEATS (DOMAIN  
 CC III).  
 CC FT 784 831 LAMININ EGF-LIKE 6.  
 CC FT 832 877 LAMININ EGF-LIKE 7.  
 CC FT 878 927 LAMININ EGF-LIKE 8.  
 CC FT 928 986 LAMININ EGF-LIKE 9.  
 CC FT 987 1038 LAMININ EGF-LIKE 10.  
 CC FT 1039 1095 LAMININ EGF-LIKE 11.  
 CC FT 1096 1143 LAMININ EGF-LIKE 12.  
 CC FT 1144 1190 LAMININ EGF-LIKE 13.  
 CC FT 1191 1410 DOMAIN II.  
 CC FT 1411 1443 DOMAIN ALPHA.  
 CC FT 1444 1799 DOMAIN I.  
 CC FT 1799 1934 COILED COIL (POTENTIAL).  
 CC FT 1935 2004 COILED COIL (POTENTIAL).  
 CC FT 2005 2074 COILED COIL (POTENTIAL).  
 CC FT 2075 2144 COILED COIL (POTENTIAL).  
 CC FT 2145 2214 BY SIMILARITY.  
 CC FT 2215 2284 BY SIMILARITY.  
 CC FT 2285 2354 BY SIMILARITY.  
 CC FT 2355 2424 BY SIMILARITY.  
 CC FT 2425 2494 BY SIMILARITY.  
 CC FT 2495 2564 BY SIMILARITY.  
 CC FT 2565 2634 BY SIMILARITY.  
 CC FT 2635 2704 BY SIMILARITY.  
 CC FT 2705 2774 BY SIMILARITY.  
 CC FT 2775 2844 BY SIMILARITY.  
 CC FT 2845 2914 BY SIMILARITY.  
 CC FT 2915 2984 BY SIMILARITY.  
 CC FT 2985 3054 BY SIMILARITY.  
 CC FT 3055 3124 BY SIMILARITY.  
 CC FT 3125 3194 BY SIMILARITY.  
 CC FT 3195 3264 BY SIMILARITY.  
 CC FT 3265 3334 BY SIMILARITY.  
 CC FT 3335 3404 BY SIMILARITY.  
 CC FT 3405 3474 BY SIMILARITY.  
 CC FT 3475 3544 BY SIMILARITY.  
 CC FT 3545 3614 BY SIMILARITY.  
 CC FT 3615 3684 BY SIMILARITY.  
 CC FT 3685 3754 BY SIMILARITY.  
 CC FT 3755 3824 BY SIMILARITY.  
 CC FT 3825 3894 BY SIMILARITY.  
 CC FT 3895 3964 BY SIMILARITY.  
 CC FT 3965 4034 BY SIMILARITY.  
 CC FT 4035 4104 BY SIMILARITY.  
 CC FT 4105 4174 BY SIMILARITY.  
 CC FT 4175 4244 BY SIMILARITY.  
 CC FT 4245 4314 BY SIMILARITY.  
 CC FT 4315 4384 BY SIMILARITY.  
 CC FT 4385 4454 BY SIMILARITY.  
 CC FT 4455 4524 BY SIMILARITY.  
 CC FT 4525 4594 BY SIMILARITY.  
 CC FT 4595 4664 BY SIMILARITY.  
 CC FT 4665 4734 BY SIMILARITY.  
 CC FT 4735 4804 BY SIMILARITY.  
 CC FT 4805 4874 BY SIMILARITY.  
 CC FT 4875 4944 BY SIMILARITY.  
 CC FT 4945 5014 BY SIMILARITY.  
 CC FT 5015 5084 BY SIMILARITY.  
 CC FT 5085 5154 BY SIMILARITY.  
 CC FT 5155 5224 BY SIMILARITY.  
 CC FT 5225 5294 BY SIMILARITY.  
 CC FT 5295 5364 BY SIMILARITY.  
 CC FT 5365 5434 BY SIMILARITY.  
 CC FT 5435 5504 BY SIMILARITY.  
 CC FT 5505 5574 BY SIMILARITY.  
 CC FT 5575 5644 BY SIMILARITY.  
 CC FT 5645 5714 BY SIMILARITY.  
 CC FT 5715 5784 BY SIMILARITY.  
 CC FT 5785 5854 BY SIMILARITY.  
 CC FT 5855 5924 BY SIMILARITY.  
 CC FT 5925 5994 BY SIMILARITY.  
 CC FT 5995 6064 BY SIMILARITY.  
 CC FT 6065 6134 BY SIMILARITY.  
 CC FT 6135 6204 BY SIMILARITY.  
 CC FT 6205 6274 BY SIMILARITY.  
 CC FT 6275 6344 BY SIMILARITY.  
 CC FT 6345 6414 BY SIMILARITY.  
 CC FT 6415 6484 BY SIMILARITY.  
 CC FT 6485 6554 BY SIMILARITY.  
 CC FT 6555 6624 BY SIMILARITY.  
 CC FT 6625 6694 BY SIMILARITY.  
 CC FT 6695 6764 BY SIMILARITY.  
 CC FT 6765 6834 BY SIMILARITY.  
 CC FT 6835 6904 BY SIMILARITY.  
 CC FT 6905 6974 BY SIMILARITY.  
 CC FT 6975 7044 BY SIMILARITY.  
 CC FT 7045 7114 BY SIMILARITY.  
 CC FT 7115 7184 BY SIMILARITY.  
 CC FT 7185 7254 BY SIMILARITY.  
 CC FT 7255 7324 BY SIMILARITY.  
 CC FT 7325 7394 BY SIMILARITY.  
 CC FT 7395 7464 BY SIMILARITY.  
 CC FT 7465 7534 BY SIMILARITY.  
 CC FT 7535 7604 BY SIMILARITY.  
 CC FT 7605 7674 BY SIMILARITY.  
 CC FT 7675 7744 BY SIMILARITY.  
 CC FT 7745 7814 BY SIMILARITY.  
 CC FT 7815 7884 BY SIMILARITY.  
 CC FT 7885 7954 BY SIMILARITY.  
 CC FT 7955 8024 BY SIMILARITY.  
 CC FT 8025 8094 BY SIMILARITY.  
 CC FT 8095 8164 BY SIMILARITY.  
 CC FT 8165 8234 BY SIMILARITY.  
 CC FT 8235 8304 BY SIMILARITY.  
 CC FT 8305 8374 BY SIMILARITY.  
 CC FT 8375 8444 BY SIMILARITY.  
 CC FT 8445 8514 BY SIMILARITY.  
 CC FT 8515 8584 BY SIMILARITY.  
 CC FT 8585 8654 BY SIMILARITY.  
 CC FT 8655 8724 BY SIMILARITY.  
 CC FT 8725 8794 BY SIMILARITY.  
 CC FT 8795 8864 BY SIMILARITY.  
 CC FT 8865 8934 BY SIMILARITY.  
 CC FT 8935 9004 BY SIMILARITY.  
 CC FT 9005 9074 BY SIMILARITY.  
 CC FT 9075 9144 BY SIMILARITY.  
 CC FT 9145 9214 BY SIMILARITY.  
 CC FT 9215 9284 BY SIMILARITY.  
 CC FT 9285 9354 BY SIMILARITY.  
 CC FT 9355 9424 BY SIMILARITY.  
 CC FT 9425 9494 BY SIMILARITY.  
 CC FT 9495 9564 BY SIMILARITY.  
 CC FT 9565 9634 BY SIMILARITY.  
 CC FT 9635 9704 BY SIMILARITY.  
 CC FT 9705 9774 BY SIMILARITY.  
 CC FT 9775 9844 BY SIMILARITY.  
 CC FT 9845 9914 BY SIMILARITY.  
 CC FT 9915 9984 BY SIMILARITY.  
 CC FT 9985 10054 BY SIMILARITY.  
 CC FT 10055 10124 BY SIMILARITY.  
 CC FT 10125 10194 BY SIMILARITY.  
 CC FT 10195 10264 BY SIMILARITY.  
 CC FT 10265 10334 BY SIMILARITY.  
 CC FT 10335 10404 BY SIMILARITY.  
 CC FT 10405 10474 BY SIMILARITY.  
 CC FT 10475 10544 BY SIMILARITY.  
 CC FT 10545 10614 BY SIMILARITY.  
 CC FT 10615 10684 BY SIMILARITY.  
 CC FT 10685 10754 BY SIMILARITY.  
 CC FT 10755 10824 BY SIMILARITY.  
 CC FT 10825 10894 BY SIMILARITY.  
 CC FT 10895 10964 BY SIMILARITY.  
 CC FT 10965 11034 BY SIMILARITY.  
 CC FT 11035 11104 BY SIMILARITY.  
 CC FT 11105 11174 BY SIMILARITY.  
 CC FT 11175 11244 BY SIMILARITY.  
 CC FT 11245 11314 BY SIMILARITY.  
 CC FT 11315 11384 BY SIMILARITY.  
 CC FT 11385 11454 BY SIMILARITY.  
 CC FT 11455 11524 BY SIMILARITY.  
 CC FT 11525 11594 BY SIMILARITY.  
 CC FT 11595 11664 BY SIMILARITY.  
 CC FT 11665 11734 BY SIMILARITY.  
 CC FT 11735 11804 BY SIMILARITY.  
 CC FT 11805 11874 BY SIMILARITY.  
 CC FT 11875 11944 BY SIMILARITY.  
 CC FT 11945 12014 BY SIMILARITY.  
 CC FT 12015 12084 BY SIMILARITY.  
 CC FT 12085 12154 BY SIMILARITY.  
 CC FT 12155 12224 BY SIMILARITY.  
 CC FT 12225 12294 BY SIMILARITY.  
 CC FT 12295 12364 BY SIMILARITY.  
 CC FT 12365 12434 BY SIMILARITY.  
 CC FT 12435 12504 BY SIMILARITY.  
 CC FT 12505 12574 BY SIMILARITY.  
 CC FT 12575 12644 BY SIMILARITY.  
 CC FT 12645 12714 BY SIMILARITY.  
 CC FT 12715 12784 BY SIMILARITY.  
 CC FT 12785 12854 BY SIMILARITY.  
 CC FT 12855 12924 BY SIMILARITY.  
 CC FT 12925 12994 BY SIMILARITY.  
 CC FT 12995 13064 BY SIMILARITY.  
 CC FT 13065 13134 BY SIMILARITY.  
 CC FT 13135 13204 BY SIMILARITY.  
 CC FT 13205 13274 BY SIMILARITY.  
 CC FT 13275 13344 BY SIMILARITY.  
 CC FT 13345 13414 BY SIMILARITY.  
 CC FT 13415 13484 BY SIMILARITY.  
 CC FT 13485 13554 BY SIMILARITY.  
 CC FT 13555 13624 BY SIMILARITY.  
 CC FT 13625 13694 BY SIMILARITY.  
 CC FT 13695 13764 BY SIMILARITY.  
 CC FT 13765 13834 BY SIMILARITY.  
 CC FT 13835 13904 BY SIMILARITY.  
 CC FT 13905 13974 BY SIMILARITY.  
 CC FT 13975 14044 BY SIMILARITY.  
 CC FT 14045 14114 BY SIMILARITY.  
 CC FT 14115 14184 BY SIMILARITY.  
 CC FT 14185 14254 BY SIMILARITY.  
 CC FT 14255 14324 BY SIMILARITY.  
 CC FT 14325 14394 BY SIMILARITY.  
 CC FT 14395 14464 BY SIMILARITY.  
 CC FT 14465 14534 BY SIMILARITY.  
 CC FT 14535 14604 BY SIMILARITY.  
 CC FT 14605 14674 BY SIMILARITY.  
 CC FT 14675 14744 BY SIMILARITY.  
 CC FT 14745 14814 BY SIMILARITY.  
 CC FT 14815 14884 BY SIMILARITY.  
 CC FT 14885 14954 BY SIMILARITY.  
 CC FT 14955 15024 BY SIMILARITY.  
 CC FT 15025 15094 BY SIMILARITY.  
 CC FT 15095 15164 BY SIMILARITY.  
 CC FT 15165 15234 BY SIMILARITY.  
 CC FT 15235 15304 BY SIMILARITY.  
 CC FT 15305 15374 BY SIMILARITY.  
 CC FT 15375 15444 BY SIMILARITY.  
 CC FT 15445 15514 BY SIMILARITY.  
 CC FT 15515 15584 BY SIMILARITY.  
 CC FT 15585 15654 BY SIMILARITY.  
 CC FT 15655 15724 BY SIMILARITY.  
 CC FT 15725 15794 BY SIMILARITY.  
 CC FT 15795 15864 BY SIMILARITY.  
 CC FT 15865 15934 BY SIMILARITY.  
 CC FT 15935 16004 BY SIMILARITY.  
 CC FT 16005 16074 BY SIMILARITY.  
 CC FT 16075 16144 BY SIMILARITY.  
 CC FT 16145 16214 BY SIMILARITY.  
 CC FT 16215 16284 BY SIMILARITY.  
 CC FT 16285 16354 BY SIMILARITY.  
 CC FT 16355 16424 BY SIMILARITY.  
 CC FT 16425 16494 BY SIMILARITY.  
 CC FT 16495 16564 BY SIMILARITY.  
 CC FT 16565 16634 BY SIMILARITY.  
 CC FT 16635 16704 BY SIMILARITY.  
 CC FT 16705 16774 BY SIMILARITY.  
 CC FT 16775 16844 BY SIMILARITY.  
 CC FT 16845 16914 BY SIMILARITY.  
 CC FT 16915 16984 BY SIMILARITY.  
 CC FT 16985 17054 BY SIMILARITY.  
 CC FT 17055 17124 BY SIMILARITY.  
 CC FT 17125 17194 BY SIMILARITY.  
 CC FT 17195 17264 BY SIMILARITY.  
 CC FT 17265 17334 BY SIMILARITY.  
 CC FT 17335 17404 BY SIMILARITY.  
 CC FT 17405 17474 BY SIMILARITY.  
 CC FT 17475 17544 BY SIMILARITY.  
 CC FT 17545 17614 BY SIMILARITY.  
 CC FT 17615 17684 BY SIMILARITY.  
 CC FT 17685 17754 BY SIMILARITY.  
 CC FT 17755 17824 BY SIMILARITY.  
 CC FT 17825 17894 BY SIMILARITY.  
 CC FT 17895 17964 BY SIMILARITY.  
 CC FT 17965 18034 BY SIMILARITY.  
 CC FT 18035 18104 BY SIMILARITY.  
 CC FT 18105 18174 BY SIMILARITY.  
 CC FT 18175 18244 BY SIMILARITY.  
 CC FT 18245 18314 BY SIMILARITY.  
 CC FT 18315 18384 BY SIMILARITY.  
 CC FT 18385 18454 BY SIMILARITY.  
 CC FT 18455 18524 BY SIMILARITY.  
 CC FT 18525 18594 BY SIMILARITY.  
 CC FT 18595 18664 BY SIMILARITY.  
 CC FT 18665 18734 BY SIMILARITY.  
 CC FT 18735 18804 BY SIMILARITY.  
 CC FT 18805 18874 BY SIMILARITY.  
 CC FT 18875 18944 BY SIMILARITY.  
 CC FT 18945 19014 BY SIMILARITY.  
 CC FT 19015 19084 BY SIMILARITY.  
 CC FT 19085 19154 BY SIMILARITY.  
 CC FT 19155 19224 BY SIMILARITY.  
 CC FT 19225 19294 BY SIMILARITY.  
 CC FT 19295 19364 BY SIMILARITY.  
 CC FT 19365 19434 BY SIMILARITY.  
 CC FT 19435 19504 BY SIMILARITY.  
 CC FT 19505 19574 BY SIMILARITY.  
 CC FT 19575 19644 BY SIMILARITY.  
 CC FT 19645 19714 BY SIMILARITY.  
 CC FT 19715 19784 BY SIMILARITY.  
 CC FT 19785 19854 BY SIMILARITY.  
 CC FT 19855 19924 BY SIMILARITY.  
 CC FT 19925 19994 BY SIMILARITY.  
 CC FT 19995 20064 BY SIMILARITY.  
 CC FT 20065 20134 BY SIMILARITY.  
 CC FT 20135 20204 BY SIMILARITY.  
 CC FT 20205 20274 BY SIMILARITY.  
 CC FT 20275 20344 BY SIMILARITY.  
 CC FT 20345 20414 BY SIMILARITY.  
 CC FT 20415 20484 BY SIMILARITY.  
 CC FT 20485 20554 BY SIMILARITY.  
 CC FT 20555 20624 BY SIMILARITY.  
 CC FT 20625 20694 BY SIMILARITY.  
 CC FT 20695 20764 BY SIMILARITY.  
 CC FT 20765 20834 BY SIMILARITY.  
 CC FT 20835 20904 BY SIMILARITY.  
 CC FT 20905 20974 BY SIMILARITY.  
 CC FT 20975 21044 BY SIMILARITY.  
 CC FT 21045 21114 BY SIMILARITY.  
 CC FT 21115 21184 BY SIMILARITY.  
 CC FT 21185 21254 BY SIMILARITY.  
 CC FT 21255 21324 BY SIMILARITY.  
 CC FT 21325 21394 BY SIMILARITY.  
 CC FT 21395 21464 BY SIMILARITY.  
 CC FT 21465 21534 BY SIMILARITY.  
 CC FT 21535 21604 BY SIMILARITY.  
 CC FT 21605 21674 BY SIMILARITY.  
 CC FT 21675 21744 BY SIMILARITY.  
 CC FT 21745 21814 BY SIMILARITY.  
 CC FT 21815 21884 BY SIMILARITY.  
 CC FT 21885 21954 BY SIMILARITY.  
 CC FT 21955 22024 BY SIMILARITY.  
 CC FT 22025 22094 BY SIMILARITY.  
 CC FT 22095 22164 BY SIMILARITY.  
 CC FT 22165 22234 BY SIMILARITY.  
 CC FT 22235 22304 BY SIMILARITY.  
 CC FT 22305 22374 BY SIMILARITY.  
 CC FT 22375 22444 BY SIMILARITY.  
 CC FT 22445 22514 BY SIMILARITY.  
 CC FT 22515 22584 BY SIMILARITY.  
 CC FT 22585 22654 BY SIMILARITY.  
 CC FT 22655 22724 BY SIMILARITY.  
 CC FT 22725 22794 BY SIMILARITY.  
 CC FT 22795 22864 BY SIMILARITY.  
 CC FT 22865 22934 BY SIMILARITY.  
 CC FT 22935 23004 BY SIMILARITY.  
 CC FT 23005 23074 BY SIMILARITY.  
 CC FT 23075 23144 BY SIMILARITY.  
 CC FT 23145 23214 BY SIMILARITY.  
 CC FT 23215 23284 BY SIMILARITY.  
 CC FT 23285 23354 BY SIMILARITY.  
 CC FT 23355 23424 BY SIMILARITY.  
 CC FT 23425 23494 BY SIMILARITY.  
 CC FT 23495 23564 BY SIMILARITY.  
 CC FT 23565 23634 BY SIMILARITY.  
 CC FT 23635 23704 BY SIMILARITY.  
 CC FT 23705 23774 BY SIMILARITY.  
 CC FT 23775 23844 BY SIMILARITY.  
 CC FT 23845 23914 BY SIMILARITY.  
 CC FT 23915 23984 BY SIMILARITY.  
 CC FT 23985 24054 BY SIMILARITY.  
 CC FT 24055 24124 BY SIMILARITY.  
 CC FT 24125 24194 BY SIMILARITY.  
 CC FT 24195 24264 BY SIMILARITY.  
 CC FT 24265 24334 BY SIMILARITY.  
 CC FT 24335 24404 BY SIMILARITY.  
 CC FT 24405 24474 BY SIMILARITY.  
 CC FT 24475 24544 BY SIMILARITY.  
 CC FT 24545 24614 BY SIMILARITY.  
 CC FT 24615 24684 BY SIMILARITY.  
 CC FT 24685 24754 BY SIMILARITY.  
 CC FT 24755 24824 BY SIMILARITY.  
 CC FT 24825 24894 BY SIMILARITY.  
 CC FT 24895 24964 BY SIMILARITY.  
 CC FT 24965 25034 BY SIMILARITY.  
 CC FT 25035 25104 BY SIMILARITY.  
 CC FT 25105 25174 BY SIMILARITY.  
 CC FT 25175 25244 BY SIMILARITY.  
 CC FT 25245 25314 BY SIMILARITY.  
 CC FT 25315 25384 BY SIMILARITY.  
 CC FT 25385 25454 BY SIMILARITY.  
 CC FT 25455 25524 BY SIMILARITY.  
 CC FT 25525 25594 BY SIMILARITY.  
 CC FT 25595 25664 BY SIMILARITY.  
 CC FT 25665 25734 BY SIMILARITY.  
 CC FT 25735 25804 BY SIMILARITY.  
 CC FT 25805 25874 BY SIMILARITY.  
 CC FT 25875 25944 BY SIMILARITY.  
 CC FT 25945 26014 BY SIMILARITY.  
 CC FT 26015 26084 BY SIMILARITY.  
 CC FT 26085 26154 BY SIMILARITY.  
 CC FT 26155 26224 BY SIMILARITY.  
 CC FT 26225 26294 BY SIMILARITY.  
 CC FT 26295 26364 BY SIMILARITY.  
 CC FT 26365 26434 BY SIMILARITY.  
 CC FT 26435 26504 BY SIMILARITY.  
 CC FT 26505 26574 BY SIMILARITY.  
 CC FT 26575 26644 BY SIMILARITY.  
 CC FT 26645 26714 BY SIMILARITY.  
 CC FT 26715 26784 BY SIMILARITY.  
 CC FT 26785 26854 BY SIMILARITY.  
 CC FT 26855 26924 BY SIMILARITY.  
 CC FT 26925 26994 BY SIMILARITY.  
 CC FT 26995 27064 BY SIMILARITY.  
 CC FT 27065 27134 BY SIMILARITY.  
 CC FT 27135 27204 BY SIMILARITY.  
 CC FT 27205 27274 BY SIMILARITY.  
 CC FT 27275 27344 BY SIMILARITY.  
 CC FT 27345 27414 BY SIMILARITY.  
 CC FT 27415 27484 BY SIMILARITY.  
 CC FT 27485 27554 BY SIMILARITY.  
 CC FT 27555 27624 BY SIMILARITY.  
 CC FT 27625 27694 BY SIMILARITY.  
 CC FT 27695 27764 BY SIMILARITY.  
 CC FT 27765 27834 BY SIMILARITY.  
 CC FT 27835 27904 BY SIMILARITY.  
 CC FT 27905 27974 BY SIMILARITY.  
 CC FT 27975 28044 BY SIMILARITY.  
 CC FT 28045 28114 BY SIMILARITY.  
 CC FT 28115 28184 BY SIMILARITY.  
 CC FT 28185 28254 BY SIMILARITY.  
 CC FT 28255 28324 BY SIMILARITY.  
 CC FT 28325 28394 BY SIMILARITY.  
 CC FT 28395 28464 BY SIMILARITY.  
 CC FT 28465 28534 BY SIMILARITY.  
 CC FT 28535 28604 BY SIMILARITY.  
 CC FT 28605 28674 BY SIMILARITY.  
 CC FT 28675 28744 BY SIMILARITY.  
 CC FT 28745 28814 BY SIMILARITY.  
 CC FT 28815 28884 BY SIMILARITY.  
 CC FT 28885 28954 BY SIMILARITY.  
 CC FT 28955 29024 BY SIMILARITY.  
 CC FT 29025 29094 BY SIMILARITY.  
 CC FT 29095 29164 BY SIMILARITY.  
 CC FT 29165 29234 BY SIMILARITY.  
 CC FT 29235 29304 BY SIMILARITY.  
 CC FT 29305 29374 BY SIMILARITY.  
 CC FT 29375 29444 BY SIMILARITY.  
 CC FT 29445 29514 BY SIMILARITY.  
 CC FT 29515 29584 BY SIMILARITY.  
 CC FT 29585 29654 BY SIMILARITY.  
 CC FT 29655 29724 BY SIMILARITY.  
 CC FT 29725 29794 BY SIMILARITY.  
 CC FT 29795 29864 BY SIMILARITY.  
 CC FT 29865 29934 BY SIMILARITY.  
 CC FT 29935 30004 BY SIMILARITY.  
 CC FT 30005 30074 BY SIMILARITY.  
 CC FT 30075 30144 BY SIMILARITY.  
 CC FT 30145 30214 BY SIMILARITY.  
 CC FT 30215 30284 BY SIMILARITY.  
 CC FT 30285 30354 BY SIMILARITY.  
 CC FT 30355 30424 BY SIMILARITY.  
 CC FT 30425 30494 BY SIMILARITY.  
 CC FT 30495 30564 BY SIMILARITY.  
 CC FT 30565 30634 BY SIMILARITY.  
 CC FT 30635 30704 BY SIMILARITY.  
 CC FT 30705 30774 BY SIMILARITY.  
 CC FT 30775 30844 BY SIMILARITY.  
 CC FT 30845 30914 BY SIMILARITY.  
 CC FT 30915 30984 BY SIMILARITY.  
 CC FT 30985 31054 BY SIMILARITY.  
 CC FT 31055 31124 BY SIMILARITY.  
 CC FT 31125 31194 BY SIMILARITY.  
 CC FT 31195 31264 BY SIMILARITY.  
 CC FT 31265 31334 BY SIMILARITY.  
 CC FT 31335 31404 BY SIMILARITY.  
 CC FT 31405 31474 BY SIMILARITY.  
 CC FT 31475 31544 BY SIMILARITY.  
 CC FT 31545 31614 BY SIMILARITY.  
 CC FT 31615 31684 BY SIMILARITY.  
 CC FT 31685 31754 BY SIMILARITY.  
 CC FT 31755 31824 BY SIMILARITY.  
 CC FT 31825 31894 BY SIMILARITY.  
 CC FT 31895 31964 BY SIMILARITY.  
 CC FT 31965 32034 BY SIMILARITY.  
 CC FT 32035 32104 BY SIMILARITY.  
 CC FT 32105 32174 BY SIMILARITY.  
 CC FT 32175 32244 BY SIMILARITY.  
 CC FT 32245 32314 BY SIMILARITY.  
 CC FT 32315 32384 BY SIMILARITY.  
 CC FT 32385 32454 BY SIMILARITY.  
 CC FT 32455 32524 BY SIMILARITY.  
 CC FT 32525 32594 BY SIMILARITY.  
 CC FT 32595 32664 BY SIMILARITY.  
 CC FT 32665 32734 BY SIMILARITY.  
 CC FT 32735 32804 BY SIMILARITY.  
 CC FT 32805 32874 BY SIMILARITY.  
 CC FT 32875 32944 BY SIMILARITY.  
 CC FT 32945 33014 BY SIMILARITY.  
 CC FT 33015 33084 BY SIMILARITY.  
 CC FT 33085 33154 BY SIMILARITY.  
 CC FT 33155 33224 BY SIMILARITY.  
 CC FT 33225 33294 BY SIMILARITY.  
 CC FT 33295 33364 BY SIMILARITY.  
 CC FT 33365 33434 BY SIMILARITY.  
 CC FT 33435 33504 BY SIMILARITY.  
 CC FT 33505 33574 BY SIMILARITY.  
 CC FT 33575 33644 BY SIMILARITY.  
 CC FT 33645 33714 BY SIMILARITY.  
 CC FT 33715 33784 BY SIMILARITY.  
 CC FT 33785 33854 BY SIMILARITY.  
 CC FT 33855 33924 BY SIMILARITY.  
 CC FT 33925 33994 BY SIMILARITY.  
 CC FT 33995 34064 BY SIMILARITY.  
 CC FT 340

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92268130.
RA Olah M.E., Ren H., Ostrowski J., Jacobson K., Stiles G.L.;
RT "Cloning, expression, and characterization of the unique bovine A1
RT site-directed mutagenesis."
RL J. Biol. Chem. 267:10764-10770(1992).
CC -1- FUNCTION: RECEPTOR FOR ADENOSINE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY G PROTEIN WHICH INHIBIT ADENYLATE CYCLASE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: X63592; CAA45135.1;
PIR: S22213; S22213.
DR HSP: P29274; 1MH.
DR GCRDB: GCR_0284; -.
DR PFAM: PF00001; 7tm1.1;
DR PRINTS: PR00237; GPCRHDOPS.
DR PRINTS: PR00424; ADENOSTINER.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR.1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 10 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 11 33 1 (POTENTIAL).
FT DOMAIN 34 46 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 47 69 2 (POTENTIAL).
FT DOMAIN 70 80 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 81 102 3 (POTENTIAL).
FT DOMAIN 103 123 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 124 146 4 (POTENTIAL).
FT DOMAIN 147 176 5 (POTENTIAL).
FT TRANSMEM 177 201 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 202 235 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 236 259 6 (POTENTIAL).
FT DOMAIN 260 267 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 268 292 7 (POTENTIAL).
FT DOMAIN 293 326 CYTOPLASMIC (POTENTIAL).
FT DISULFID 80 169 BY SIMILARITY.
FT CARBOHYD 148 148 POTENTIAL.
FT CARBOHYD 159 309 PALMITATE (POTENTIAL).
FT LIPID 101 101 M -> I (IN REF. 2).
FT CONFLICT 101 101
SQ SEQUENCE 326 AA; 36597 MW; 1697A5F03A57285 CRC64;

Query Match 5.5%; Score 96; DB 1; Length 326;
Best Local Similarity 37.1%; Pred. No. 1.27e-01;
Matches 13; Conservative 11; Mismatches 8; Indels 3; Gaps 3;

```

```

GN CLU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92184774.
RA Diemer V., Hoyle M., Baglioni C., Mills A.J.;
RT "Expression of porcine complement cytotoxicity inhibitor mRNA in
RT cultured aortic smooth muscle cells. Changes during differentiation
RT in vitro."
RL J. Biol. Chem. 267:5257-5264(1992).
CC -1- FUNCTION: NOT YET CLEAR. IT IS KNOWN TO BE EXPRESSED IN A VARIETY
CC OF TISSUES AND IT SEEMS TO BE ABLE TO BIND TO CELLS, MEMBRANES,
CC AND HYDROPHOBIC PROTEINS. IT HAS BEEN ASSOCIATED WITH PROGRAMMED
CC CELL DEATH (BY SIMILARITY).
CC -1- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED HETERODIMER (BY
CC SIMILARITY).
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN BRAIN AND LIVER; LOWER
CC LEVELS ARE DETECTED IN OTHER TISSUES, INCLUDING THE AORTA.
CC -1- SIMILARITY: BELONGS TO THE CLUSTERIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M84639; AAA1013.1;
DR PFAM: PF01093; Clusterin.1.
DR PROSITE: PS00493; CLUSTERIN_1;
KW Glycoprotein; Signal.
FT SIGNAL 1 28
FT CHAIN 29 446 BY SIMILARITY.
FT CHAIN 29 446 CLUSTERIN.
FT CHAIN 29 446 BETA-CHAIN (A CHAIN).
FT CHAIN 228 446 ALPHA-CHAIN (B CHAIN).
FT DISULFID 102 312 INTERCHAIN (BY SIMILARITY).
FT DISULFID 113 304 INTERCHAIN (BY SIMILARITY).
FT DISULFID 116 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 121 294 INTERCHAIN (BY SIMILARITY).
FT DISULFID 129 284 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 86 86 POTENTIAL.
FT CARBOHYD 103 103 POTENTIAL.
FT CARBOHYD 145 145 POTENTIAL.
FT CARBOHYD 290 290 POTENTIAL.
FT CARBOHYD 316 316 POTENTIAL.
FT CARBOHYD 353 353 POTENTIAL.
FT CARBOHYD 373 373 POTENTIAL.
SQ SEQUENCE 446 AA; 51774 MW; BID5B43466B83AA CRC64;

Query Match 5.5%; Score 96; DB 1; Length 446;
Best Local Similarity 23.5%; Pred. No. 1.27e-01;
Matches 8; Conservative 19; Mismatches 7; Indels 0; Gaps 0;

```

```

DB 236 LALILFALSWPLHILNCITLFCPSCHMPRII 270
OY 9 LTIATLMF-LTWLP-ESLSCAKALCAS-DYSKCLI 40

RESULT 8
ID CLOS_PIG STANDARD; PRT; 446 AA.
AC Q29549;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CLUSTERIN PRECURSOR (COMPLEMENT CYTOLYSIS INHIBITOR) (CLI).

```

```

DB 223 SRFARINMPFLPTDLYHDMFOPEDMIIHOAQ 256
OY 112 TQNNNIVSFPVAEELSHENLVSELEVNOPHH 145

RESULT 9
ID TNR2_HUMAN STANDARD; PRT; 461 AA.
AC P20333;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR
DE BINDING PROTEIN 2) (TNFR2) (P80) (TNF-R2) (P75) (CD120b) (ETANERCEPT).
GN TNFRSF1B OR TNFR2 OR TNFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 90260639.  
 RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,  
 RA Dower S.K., Cosman D., Goodwin R.G.;  
 RT "A receptor for tumor necrosis factor defines an unusual family of  
 RT cellular and viral proteins.";  
 RL Science 248:1019-1023(1990).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 91045991.  
 RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,  
 RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;  
 RT "A second tumor necrosis factor receptor gene product can shed a  
 RT naturally occurring tumor necrosis factor inhibitor.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).  
 [3]  
 [3] SEQUENCE FROM N.A.  
 RX MEDLINE: 96299745.  
 RA Bellinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,  
 RA Lepassier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,  
 RA Brodeur G.M.;  
 RT "Physical mapping and genomic structure of the human TNFR2 gene.";  
 RL Genomics 35:94-100(1996).  
 RN [4]  
 RN SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE: 90349572.  
 RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,  
 RA Ringold G.M.;  
 RT "Complementary DNA cloning of a receptor for tumor necrosis factor  
 RT and demonstration of a shed form of the receptor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).  
 [5]  
 [5] SEQUENCE OF 27-31.  
 RX MEDLINE: 90110215.  
 RA Engelmann H., Novick D., Wallach D.;  
 RT "Two tumor necrosis factor-binding proteins purified from human  
 RT urine. Evidence for immunological cross-reactivity with cell surface  
 RT tumor necrosis factor receptors.";  
 RL J. Biol. Chem. 265:1531-1536(1990).  
 RN [6]  
 RN SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.  
 RX MEDLINE: 91056048.  
 RA Loetscher H., Schlaeger E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,  
 RA Brockhaus M.;  
 RT "Purification and partial amino acid sequence analysis of two  
 RT distinct tumor necrosis factor receptors from HL60 cells.";  
 RL J. Biol. Chem. 265:20131-20138(1990).  
 RN [7]  
 RN CHARACTERIZATION.  
 RX MEDLINE: 93016040.  
 RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,  
 RA Lipari M.T., Goeddel D.V.;  
 RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.  
 RT Characterization of ligand binding, internalization, and receptor  
 RT phosphorylation.";  
 RL J. Biol. Chem. 267:21172-21178(1992).  
 RN [8]  
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.  
 RX MEDLINE: 99221490.  
 RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;  
 RT "Structural basis for self-association and receptor recognition of  
 RT human TRAF2.";  
 RL Nature 398:533-538(1999).  
 CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND  
 CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW  
 CC LEVEL ON THREONINE RESIDUES.  
 CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND  
 CC WYETH-AVENTIS). USED TO TREAT MODERATE TO SEVERE RHEUMATOID  
 CC ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING

CC PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO  
 CC TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd120b.htm".  
 CC -1- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel;  
 CC WWW="http://www.enbrelinfo.com/".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: M3315; AAA59929.1; -  
 DR EMBL: M35857; AAA63262.1; -  
 DR EMBL: U52165; AAC50622.1; JOINED.  
 DR EMBL: U52156; AAC50622.1; JOINED.  
 DR EMBL: U52157; AAC50622.1; JOINED.  
 DR EMBL: U52158; AAC50622.1; JOINED.  
 DR EMBL: U52159; AAC50622.1; JOINED.  
 DR EMBL: U52160; AAC50622.1; JOINED.  
 DR EMBL: U52161; AAC50622.1; JOINED.  
 DR EMBL: U52162; AAC50622.1; JOINED.  
 DR EMBL: U52163; AAC50622.1; JOINED.  
 DR EMBL: U52164; AAC50622.1; JOINED.  
 DR EMBL: M55994; AAA36755.1; -  
 DR PIR: A35356; A35356.  
 DR PIR: A36007; A36007.  
 DR PIR: A36475; A36475.  
 DR PIR: B35010; B35010.  
 DR PIR: A23666; A23666.  
 DR PDB: 1CA9; 12-APR-99.  
 DR MTM: 191191; -  
 DR PRAW: PF00020; TNFR\_C6; 4.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE: PS50050; TNFR\_NGFR\_2; 4.  
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;  
 KW Phosphorylation; Pharmaceutical; 3D-structure.  
 FT SIGNAL 1 22  
 FT CHAIN 23 461  
 FT DOMAIN 23 257  
 FT TRANSMEM 258 287  
 FT DOMAIN 288 461  
 FT REPEAT 39 201  
 FT REPEAT 39 201  
 FT REPEAT 77 118  
 FT REPEAT 119 162  
 FT REPEAT 163 201  
 FT DISULFID 40 53  
 FT DISULFID 54 67  
 FT DISULFID 57 75  
 FT DISULFID 78 93  
 FT DISULFID 96 110  
 FT DISULFID 100 118  
 FT DISULFID 120 126  
 FT DISULFID 134 143  
 FT DISULFID 137 161  
 FT DISULFID 164 179  
 FT CARBOHYD 171 171  
 FT CARBOHYD 193 193  
 FT CONFLICT 141 141  
 FT CONFLICT 196 196  
 FT CONFLICT 363 363  
 SQ SEQUENCE 461 AA; 48316 MW; 603B580EC67636F CRC64;  
 Query Match 5.58; Score 95; DB 1; Length 461;  
 Best Local Similarity 42.48; Pred. No. 1.80e-01;  
 Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;  
 Db 89 WPECLSCG-SRCSDDVETQACTREONRICICRPG-WICALSKQSGCRIC-APLRK-C 143



Query Match	Best Local	Similarity	Score	DB	Length
Matches 11; Conservative <td>11; Conservative <td>7; Mismatches 12; Indels 0; Gaps 0;</td> <td>5.5%;</td> <td>36.7%;</td> <td>1.27e-01;</td> </td>	11; Conservative <td>7; Mismatches 12; Indels 0; Gaps 0;</td> <td>5.5%;</td> <td>36.7%;</td> <td>1.27e-01;</td>	7; Mismatches 12; Indels 0; Gaps 0;	5.5%;	36.7%;	1.27e-01;
Db 1041	CTCNLTGTPQRCPESTDLCCHDPSNGGQPC 1070				
07	26 CNKALCASYKSKCLQELQCQCPGEGNCSC 55				
RESULT 12	STANDARD;	PRT; 3084 AA.			
AC	PI9137;				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	LAMININ ALPHA-1 CHAIN PRECURSOR (LAMININ A CHAIN).				
GN	LAM1 OR LAMA-1 OR LAMA.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
RN	[1]				

RP MEDLINE: FROM N.A., AND PARTIAL SEQUENCE.  
RA Sasaki M., Kleiman H.K., Huber H., Deutzmann R., Yamada Y.;  
RT "Laminin, a multidomain protein. The A chain has a unique globular  
RT domain and homology with the basement membrane proteoglycan and the  
RT laminin B chains";  
RL J. Biol. Chem. 263:16536-16544(1988).  
RN [2]  
RX SEQUENCE OF 1-339 FROM N.A.  
RX MEDLINE: 88225080.  
RA Hartl L., Oberhaeumer I., Deutzmann R.;  
RT "The N terminus of laminin A chain is homologous to the B chains.";  
RL Eur. J. Biochem. 173:629-635(1988).  
RN [3]  
RX SEQUENCE OF 2538-3084 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE: 89030693.  
RA Deutzmann R., Huber J., Schmetz K.A., Oberhaeumer I., Hartl L.;  
RT "Structural study of long arm fragments of laminin. Evidence for  
RT repetitive C-terminal sequences in the A-chain, not present in the B-  
RT chains";  
RL Eur. J. Biochem. 177:35-45(1988).  
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR. LAMININ  
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ) AND  
CC LAMININ-3 (S-LAMININ).  
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR  
CC COMPONENT).  
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.  
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.  
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>).  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J04064; AAA39410.1; -  
DR EMBL: X07737; CAA30561.1; -  
DR EMBL: X13459; CAA31807.1; -  
DR EMBL: M36775; AAA39406.1; -  
DR PIR: A31771; MMSA.  
DR HSSP: P02468; 1TLE.  
DR MGD: MGI:98892; LAMN1.  
DR PFAM: PF00052; laminin\_B; 2.  
DR PFAM: PF00053; laminin\_EGF; 15.  
DR PFAM: PF00054; laminin\_G; 5.  
DR PFAM: PF00055; laminin\_Nterm; 1.  
DR PRINTS: PRO0011; EGF\_LAMININ.  
DR PROSITE: PS00022; EGF\_1; 11.  
DR PROSITE: PS01186; EGF\_2; 3.  
DR PROSITE: PS01248; LAMININ\_Type\_EGF; 15.  
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
FT SIGNAL 1 24  
FT CHAIN 1 25 3084 LAMININ ALPHA-1 CHAIN.  
FT MOD RES 25 25 276 BLOCKED.  
FT DOMAIN 25 276 LAMININ N-TERMINAL (DOMAIN VI).  
FT 277 519 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN  
FT VI).











FT	DISULFID	1556	1556	INTERCHAIN (PROBABLE).
FT	DISULFID	1560	1560	INTERCHAIN (PROBABLE).
FT	CARBOHYD	38	38	POTENTIAL.
FT	CARBOHYD	555	555	POTENTIAL.
FT	CARBOHYD	665	665	POTENTIAL.
FT	CARBOHYD	763	763	POTENTIAL.
FT	CARBOHYD	926	926	POTENTIAL.
FT	CARBOHYD	952	952	POTENTIAL.
FT	CARBOHYD	1045	1045	POTENTIAL.
FT	CARBOHYD	1407	1407	POTENTIAL.
FT	CARBOHYD	1579	1579	POTENTIAL.
FT	CARBOHYD	1596	1596	POTENTIAL.
FT	CARBOHYD	1678	1678	POTENTIAL.
FT	CARBOHYD	1689	1689	POTENTIAL.
FT	CARBOHYD	1698	1698	POTENTIAL.
FT	CARBOHYD	1717	1717	POTENTIAL.
FT	CARBOHYD	1804	1804	POTENTIAL.
FT	CARBOHYD	1894	1894	POTENTIAL.
FT	CARBOHYD	1898	1898	POTENTIAL.
FT	CARBOHYD	1957	1957	POTENTIAL.
FT	CARBOHYD	1974	1974	POTENTIAL.
FT	CARBOHYD	1991	1991	POTENTIAL.
FT	CARBOHYD	2038	2038	POTENTIAL.
FT	CARBOHYD	2047	2047	POTENTIAL.
FT	CARBOHYD	2094	2094	POTENTIAL.

Note: remainder of annotations omitted.

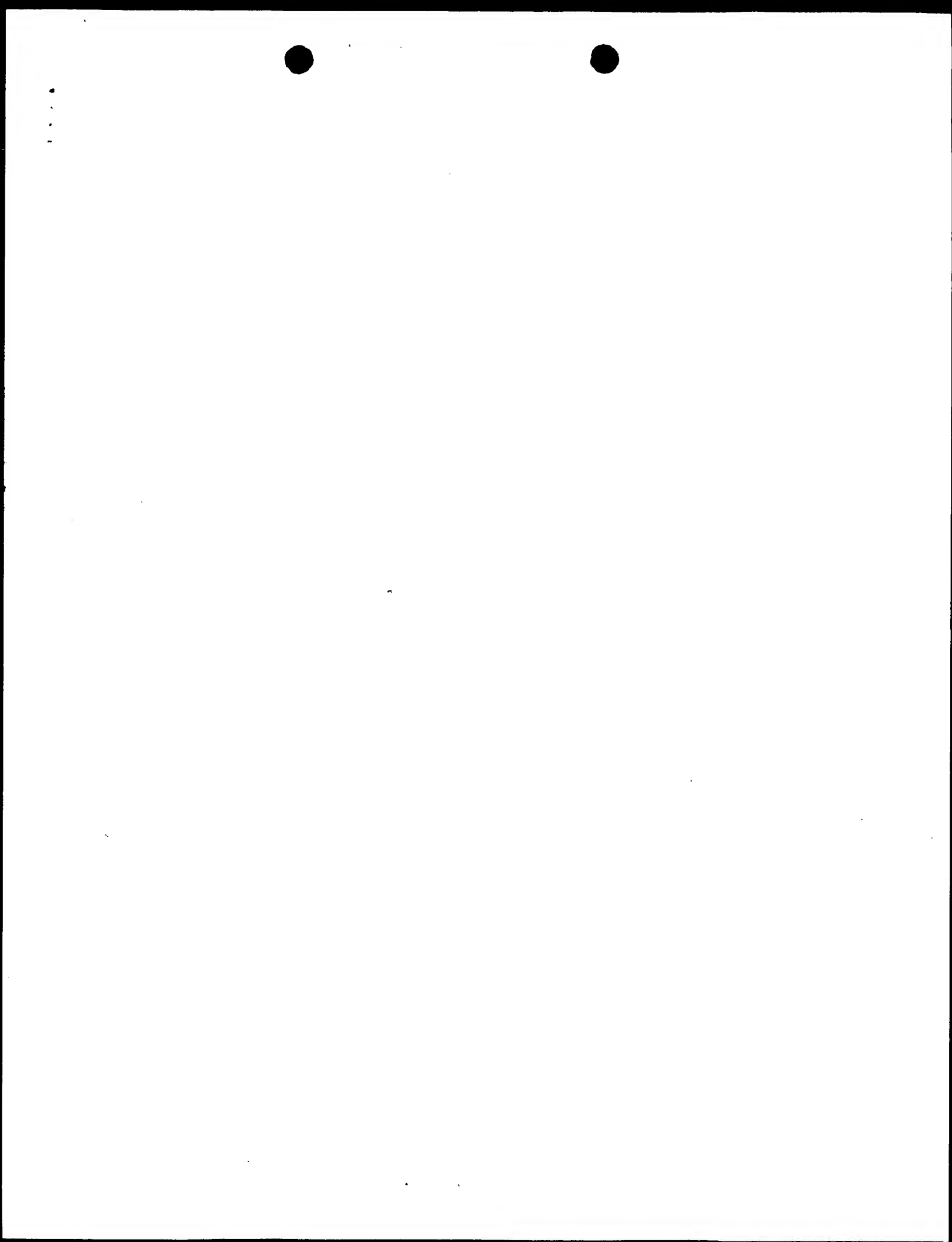
## Query Match

Best Local Similarity 5.4%; Score 94; DB 1; Length 3075;

Matches 14; Conservative 8; Mismatches 16; Indels 2; Gaps 2;

DB 1084 FPDGVPDCLDRLGTSGDAGNLEQGLGCGVEETGACPC-KE 1122  
QY 20 LPESLSCNKALCASDVSKC-LIGELCQCPRGEGNCSCCKE 58

Search completed: Wed Aug 16 09:32:22 2000  
Job time : 42 secs.



\*\*\*\*\*  
 W P S E R F (TM)  
 \*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

arch-pp protein - protein database search, using Smith-Waterman algorithm  
 Run on: Wed Aug 16 09:30:03 2000; Maspar time 13.39 Seconds  
 Tabular output not generated. 785.505 Million cell updates/sec

Title: >US-09-416-267-2  
 Description: (1-223) from US09416267.pep  
 Perfect Score: 1738  
 Sequence: 1 MKLHYAVLTALIMFLTWL.....ISPECIDGSKTVKCMGME 223

Scoring table:  
 Gap 11  
 PAM 150

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: pir64  
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 43.550; Variance 80.722; scale 0.540

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	537	30.9	249	2	A53836 developmental protein	1.06e-87
2	104	6.0	798	2	integrin beta-7 chain	4.19e-02
3	102	5.9	417	2	hypothetical protein	8.05e-02
4	101	5.8	194	2	hypothetical protein	1.11e-01
5	100	5.8	417	2	hypothetical protein	1.54e-01
6	97	5.6	244	2	glutathione transferase	3.99e-01
7	97	5.6	806	2	integrin beta-7 chain	5.45e-01
8	96	5.5	326	2	adenosine receptor A1	5.45e-01
9	96	5.5	446	2	clusterin precursor	5.45e-01
10	95	5.5	461	1	tumor necrosis factor	7.45e-01
11	95	5.5	471	2	holx protein - Rhlzob	7.45e-01
12	96	5.5	1801	1	laminin beta-2 chain	5.45e-01
13	96	5.5	3084	1	laminin alpha-1 chain	5.45e-01
14	94	5.4	62	2	metallothionein - Afl	1.01e+00
15	94	5.4	3075	2	laminin alpha-1 chain	1.01e+00
16	92	5.3	1303	2	hypothetical protein	1.87e+00
17	92	5.3	404	2	hypothetical protein	1.87e+00
18	92	5.3	627	2	membrane transport pr	1.87e+00
19	91	5.2	53	2	metallothionein - rab	3.41e+00
20	90	5.2	63	2	metallothionein 1 - p	3.41e+00
21	90	5.2	116	2	dnar-related protein	3.41e+00
22	90	5.2	157	2	probable olfactory re	3.41e+00
23	90	5.2	167	2	conserved hypothetica	3.41e+00

24	90	5.2	241	2	T16802	hypothetical protein	3.41e+00
25	90	5.2	324	2	S53502	A(1) adenosine recept	3.41e+00
26	91	5.2	445	2	A40018	clusterin precursor -	2.53e+00
27	91	5.2	618	2	S69067	probable glycine--trn	2.53e+00
28	90	5.2	690	2	S54775	cell size regulation	3.41e+00
29	90	5.2	925	1	A39216	plasma cell membrane	3.41e+00
30	88	5.1	43	1	SMFR2	metallothionein 2 - f	6.15e+00
31	88	5.1	113	1	S07092	gonadotropin beta cha	6.15e+00
32	88	5.1	140	2	A48165	gonadotropin II beta	6.15e+00
33	89	5.1	244	2	UC2425	glutathione transfera	4.59e+00
34	89	5.1	326	2	A53005	adenosine receptor A1	4.59e+00
35	89	5.1	326	2	C30341	G protein-coupled rec	4.59e+00
36	89	5.1	328	2	JN0675	adenosine receptor A1	4.59e+00
37	88	5.1	586	1	TYVPBP	large T antigen - bov	6.15e+00
38	88	5.1	757	2	T16609	hypothetical protein	6.15e+00
39	88	5.1	1042	2	S43904	hyaluronidase - Clost	6.15e+00
40	88	5.1	1652	2	I50711	complement C3 precurs	6.15e+00
41	89	5.1	1964	2	T09059	notch4 - mouse	4.59e+00
42	89	5.1	3635	2	T10053	laminin alpha 5 chain	4.59e+00
43	87	5.0	185	2	T14831	adenosine receptor su	8.24e+00
44	87	5.0	185	2	T15623	hypothetical protein	8.24e+00
45	87	5.0	565	2	T16408	hypothetical protein	8.24e+00

## ALIGNMENTS

RESULT 1  
 ENTRY A53836 #type complete  
 TITLE developmental protein TSG precursor - fruit fly (Drosophila melanogaster)  
 ORGANISM #formal\_name Drosophila melanogaster  
 DATE 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 24-Sep-1998

ACCESSIONS A53836  
 REFERENCE A53836  
 #authors Mason, E.D.; Konrad, K.D.; Webb, C.D.; Marsh, J.L.  
 #journal Genes Dev. (1994) 8:1489-1501  
 #title Dorsal midline fate in Drosophila embryos requires twisted gastrulation, a gene encoding a secreted protein related to human connective tissue growth factor.

#cross-references MIM:95047309  
 #accession A53836  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-249 #label MAS  
 #cross-references GB:U09808; NID:g3006232; PID:g529900

GENETICS  
 #gene tsf  
 #cross-references FlyBase:FBgn0003865

KEYWORDS  
 #domain signal sequence #status predicted #label SIG\

FEATURE  
 1-21 #product developmental protein tsf #status predicted  
 22-249 #label MAT

SUMMARY  
 #length 249 #molecular-weight 27220 #checksum 1494

Query Match 30.9%; Score 537; DB 2; Length 249;  
 Best local similarity 35.4%; Pred. No. 1.06e-87;  
 Matches 79; Conservative 54; Mismatches 71; Indels 19; Gaps 13;

Db	1	MOLICGVIIIFVGIAPV-SSLANDGCGNEVYGSVSKCLITGSCCKLND--CHCCKRC	57
Qy	1	MLKLH-YVAVLTALIMFLTWLPESLSCNKALCASDVSKCLOELCCRRGEGSCCKRC	59
Db	58	LNCIGELIYECGGLDMC-PKHQDVLPSTLRSELDI-EGVELEFDTLTAEDDE-GWST	114
Qy	60	MLCLGALMDCCDCVGMCPNRYSDTPP-TSKSTVELLEPLISLRALLEGDTLNMNI	118
Db	115	IRFSMRAGFKOR--V--OG-GAS--GDAGNGNGNGNG-SAGVT-LCTVIYVNSCIRA	163
Qy	119	VSFVVAELHHEHNTVLFETVNGPHHONVSPSNVHAAPYSSDKRMCTVYFDSCMSI	178
Db	164	NKRCQCCESMGASSYRMFHDCGCEVGCENCLNTGINESRCGC	206



```

ACCESSIONS      S03806
REFERENCE        S03803
#authors        Rattio, M.; Jalli, T.; Saraste, M.
#journal        EMBO J. (1987) 6:2825-2833
#title          Isolation and analysis of the genes for cytochrome c oxidase
                  in Paracoccus denitrificans.
#accession      S03806
##molecule_type DNA
##residues     1-194 ##label RAI
##cross-references EMBL:X05826; NID:g45468; PIDN:CAA29271.1; PID:g45472
CLASSIFICATION #superfamily cytochrome-c oxidase assembly protein COX11
SUMMARY         #length 194 #molecular-weight 21325 #checksum 9149

Query Match      5.8%; Score 101; DB 2; Length 194;
Best Local Similarity 25.5%; Pred. No. 1,11e-01;
Matches 25; Conservative 20; Mismatches 46; Indels 5; Gaps 5;

      23 MGALSAAPFYFSGWCFAGTGTVAAPASDVTDEKIRRF-D-ANADSNIGWTFRP 80
      :||| : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 LGAL-WPECCDCGCMCRPRYSDPPTSTSYBE-LHEPIPSLFRALRGDTQLNNMIVS 120
      : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 81 MOREMEIKGENAIAFYEAIAINNTDEPVGTASTYV-AP 117
      : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 FPAVEELSHENLVSFLETYNQPHQNVSPSNVAVAP 158
      : : : : : : : : : : : : : : : : : : : : : : : : : :

.RESULT 5
ENTRY      T08724 #type complete
TITLE      hypothetical protein DKFZp566D21.1 - human
ORGANISM   Homo sapiens #common_name man
DATE       11-Jun-1989 #sequence_revision 11-Jun-1999 #text_change
          13-Aug-1999

ACCESSIONS T08724
REFERENCE   216468
#authors    Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.;
            Wiemann, S.

#submission submitted to the Protein Sequence Database, May 1999
#accession  T08724
##molecule_type mRNA
##residues 1-417 ##label KOE
##cross-references EMBL:AL050275
##experimental_source fetal kidney; clone DKFZp566D213

GENETICS
#note       DKFZp566D213.1
SUMMARY     #length 417 #molecular-weight 44934 #checksum 2694

Query Match      5.8%; Score 100; DB 2; Length 417;
Best Local Similarity 28.6%; Pred. No. 1.54e-01;
Matches 14; Conservative 12; Mismatches 21; Indels 2; Gaps 2;

Db 236 WALHTKACVDIDECGTGAGANGADQFCVNTGSGYECRCAKACLGOMGA 284
      | : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 19 WLPESLSC-NKALCASPYSKCLIGELCQCRPGGNC-SCCKECMCLTGA 65
      : : : : : : : : : : : : : : : : : : : : : : : : : :

.RESULT 6
ENTRY      S71879 #type complete
TITLE      glutathione transferase (EC 2.5.1.18) theta 2, hepatic -
            mouse
ORGANISM   Mus musculus #common_name house mouse
DATE       21-Apr-1997 #sequence_revision 09-May-1997 #text_change
          18-Jun-1999
      S71879
      S71879
      Malwareing, G.W.; Williams, S.M.; Foster, J.R.; Tugwood, J.;
      Green, T.
      Blochem. J. (1996) 318:297-303
      The distribution of Theta-class glutathione S-transferases in
      the liver and lung of mouse, rat and human.
#cross-references MUID:96358519
#accession   S71879
#status      preliminary
#molecule_type mRNA

```

```

#residues      1-244 #label MAI
#cross-references EMBL:X98056; NID:q1340077; PIDN:CAA66666.1;
                  PID:e245091; PID:q1340078
#note          the authors translated the codon TCA for residue 11 as
                  Ala
CLASSIFICATION #superfamily glutathione transferase
KEYWORDS      transferase
SUMMARY       #length 244 #molecular-weight 27634 #checksum 5537

Query Match      5.68; Score 97; DB 2; Length 244;
Best Local Similarity 30.08; Pied. No. 3,99e-01;
Matches 21; Conservative 21; Mismatches 22; Indels 6; Gaps 6;

Db 169 SLEEMOPVALGK-NIEGRRPOLTAREAREAF-LGAEICQEAH-STIISIGQAARKKL 225
OY :||| :||| : ||| ||| : ||| : ||| : ||| : ||| : ||| : ||| :
   92 TVEELHEPISTLRALTEGDTQLN-W-NIV-SEPPAEELSHENLVSELETYNQPHQNV 148
Dd 226 PVPPEPVHAS 235
OY :||| :||| :
   149 SVSPNNVHAP 158

RESULT 7
ENTRY      A46271 #type complete
TITLE      integrin beta-7 chain precursor - mouse
ALTERNATE_NAMES M290 antigen beta chain
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
          20-Aug-1999
ACCESSIONS A46271; A42483; B46503; PN0017; A60229
REFERENCE   A46271
#authors    Hu, M.C.; Crowe, D.T.; Weissman, I.L.; Holzman, B.
#journal     Proc. Natl. Acad. Sci. U.S.A. (1992) 89:8254-8258
#title       Cloning and expression of mouse integrin beta p(beta 7): a
             functional role in Peyer's patch-specific lymphocyte
             homing.
#cross-references MUID:92390425
#accession   A46271
#status      Preliminary
#molecule_type nucleic acid
#residues    1-806 #label HUI
#note        sequence extracted from NCBI backbone (NCBIN:113097,
             NCBI:P.113125)

REFERENCE   A42483
#authors     Yuan, Q.; Jiang, W.W.; Leung, E.; Hollander, D.; Watson,
             J.D.; Kristiansen, G.W.
#journal     J. Biol. Chem. (1992) 267:7352-7358
#title       Molecular cloning of the mouse Integrin beta 7 subunit.
#cross-references MUID:92218384
#accession   A42483
#molecule_type mRNA
#residues    1-123,125-556,'H',558-806 #label YUA
#cross-references GB:G68903; NID:q349598; PIDN:AAA02749.1; PID:q349599
#note        sequence extracted from NCBI backbone (NCBIN:93455,
             NCBI:P.93463)

REFERENCE   A46503
#authors     Gurish, M.F.; Bell, A.F.; Smith, T.J.; Ducharme, L.A.; Wang,
             R.K.; Weis, J.H.
#journal     J. Immunol. (1992) 149:1964-1972
#title       Expression of murine beta 7, alpha 4, and beta 1 integrin
             genes by rodent mast cells.
#cross-references MUID:92388664
#accession   B46503
#status      preliminary
#molecule_type mRNA
#residues    1-80,'G',82-537,'C',539-806 #label GUR
#cross-references GB:S44607; NID:q255138; PIDN:AA623193.1; PID:q255139
#note        sequence extracted from NCBI backbone (NCBIN:113091,
             NCBI:P.113092)

REFERENCE   PN0017
#authors     Yuan, Q.; Jiang, W.; Hollander, D.; Leung, E.; Watson, J.D.;
             Kristiansen, G.W.

```

#Journal Biochem. Biophys. Res. Commun. (1991) 176:1443-1449  
#title Identity between the novel integrin beta 7 subunit and an antigen found highly expressed on intraepithelial lymphocytes in the small intestine.  
#cross-references MUID:91248239  
#accession F00017  
#status nucleic acid sequence not shown  
#molecule\_type mRNA  
#residues 1-80, 'E', 82-90 ##label YU2  
#journal A60229  
#authors Kishaw, P.J.; Murant, S.J.  
#journal Eur. J. Immunol. (1990) 20:2201-2207  
#title A new surface antigen on intraepithelial lymphocytes in the intestine.  
#cross-references MUID:91055592  
#accession A60229  
#molecule\_type protein  
#residues 20-32 ##label KITL  
CLASSIFICATION #superfamily integrin beta chain; laminin-type EGF-like homology  
KEYWORDS cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein; heterodimer; membrane protein; phosphoprotein  
FEATURE 1-19  
20-90 #domain signal sequence #status predicted #label SIG  
#product integrin beta 7 chain (fragment) #status experimental #label IBC  
SUMMARY #length 806 #molecular-weight 87411 #checksum 9972  
Query Match 5.6%; Score 97; DB 2; Length 806;  
Best Local Similarity 31.6%; Pred. No. 3,99e-01;  
Matches 18; Conservative 12; Mismatches 23; Indels 4; Gaps 4;  
Db 598 CSKVSVCSPGEGIC-SGHGYCKNCRCQCLDGYGALCDGCKSPCEQRPDCAE 653  
OY 31 CASDVSKCLIGELCCGRPEGNCCK-ECM-LCIGALMDECCCVGKCNF-RNYSD 84  
RESULT 8  
ENTRY A38144 #type complete  
TITLE adenosine receptor A1 - bovine  
ORGANISM #formal\_name Bos primigenius taurus #common\_name cattle  
DATE 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Sep-1999  
ACCESSION A38144; S20390; S22213  
REFERENCE A38144  
#authors Olah, M.E.; Ren, H.; Ostrowski, J.; Jacobson, K.A.; Stiles, G.L.  
#journal J. Biol. Chem. (1992) 267:10764-10770  
#title Cloning, expression, and characterization of the unique bovine A1 adenosine receptor. Studies on the ligand binding site by site-directed mutagenesis.  
#cross-references MUID:92268130  
#accession A38144  
#molecule\_type mRNA  
#residues 1-326 ##label OLA  
#cross-references GB:M6261; NID:9162597; PIDN:AAA30350.1; PID:9162598  
#experimental\_source brain  
#note Sequence extracted from NCBI backbone (NCBIN:103815, NCBIPI:103816)  
REFERENCE S20390  
#authors Tucker, A.L.; Linden, J.; Robeva, A.S.; D'Angelo, D.D.; Lynch, K.R.  
#journal FEBS Lett. (1992) 297:107-111  
#title Cloning and expression of a bovine adenosine A(1)-receptor cDNA.  
#cross-references MUID:92201360  
#accession S20390  
#status nucleic acid sequence not shown  
#molecule\_type mRNA  
#residues 1-100, 'W', 102-326 ##label TUC  
#cross-references EMBL:X63592  
#experimental\_source brain

CLASSIFICATION #superfamily adenosine receptor A1  
KEYWORDS adenylylate cyclase inhibitor; G protein-coupled receptor; glycoprotein; transmembrane protein  
FEATURE 148,159  
#binding\_site carbohydrate (Asn) (covalent) #status predicted  
SUMMARY #length 326 #molecular-weight 36579 #checksum 6464  
Query Match 5.5%; Score 96; DB 2; Length 326;  
Best Local Similarity 37.1%; Pred. No. 5.45e-01;  
Matches 13; Conservative 11; Mismatches 8; Indels 3; Gaps 3;  
Db 236 LALILFALSWLPHILNCITRCPSCMPEIL 270  
OY 9 LTLAILMF-LTWLP-BLSLNCNALCAS-DVSKCLI 40

RESULT 9  
ENTRY A42108 #type complete  
TITLE clusterin precursor - pig  
ALTERNATE\_NAMES complement cytolysis inhibitor; CP40 protein  
ORGANISM #formal\_name Sus scrofa domestica #common\_name domestic pig  
DATE 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
ACCESSION A42108; JC5535; PC4475  
REFERENCE A42108  
#authors Diemer, V.; Hoyle, M.; Baglioni, C.; Millis, A.J.  
#journal J. Biol. Chem. (1992) 267:5257-5264  
#title Expression of porcine complement cytolysis inhibitor mRNA in cultured aortic smooth muscle cells. Changes during differentiation in vitro.  
#cross-references MUID:92184774  
#accession A42108  
#molecule\_type mRNA  
#residues 1-446 ##label DTE  
#cross-references GB:M84639; NID:9164408; PIDN:AAA31013.1; PID:9164409  
#experimental\_source aortic smooth muscle cells  
#note sequence extracted from NCBI backbone (NCBIN:87354, NCBIPI:87356)  
REFERENCE JC5535  
#authors Ogawa, S.; Ishibashi, Y.; Sakamoto, Y.; Kitamura, K.; Kubo, M.; Sakai, T.; Inoue, K.  
#journal Biochem. Biophys. Res. Commun. (1997) 234:712-718  
#title The glycoproteins that occur in the colloids of senescent porcine pituitary glands are clusterin and glycosylated albumin fragments.  
#cross-references MUID:97318844  
#accession JC5535  
#molecule\_type DNA  
#residues 1-446 ##label OGA  
#accession PC4475  
#molecule\_type protein  
#residues 58-66:68-77; 229-247; 249-251; 408-436 ##label OGT  
#experimental\_source pituitary gland  
CLASSIFICATION #superfamily clusterin  
FEATURE 23-227  
23-227 #domain signal sequence #status predicted #label SIG  
228-446 #domain clusterin beta chain #status predicted #label BCN  
SUMMARY #length 446 #molecular-weight 51774 #checksum 562  
Query Match 5.5%; Score 96; DB 2; Length 446;  
Best Local Similarity 23.5%; Pred. No. 5.45e-01;  
Matches 8; Conservative 19; Mismatches 7; Indels 0; Gaps 0;  
Db 223 SRFANIMPFLFDLNYHDFOPFFDMHQAOQ 256  
OY 112 TOLMNIYSPVAVELSHENLYVFLFTVNP 145



```

RESULT 10
ENTRY #type complete
A35356 tumor necrosis factor receptor type 2 precursor - human
TITLE 75K tumor necrosis factor receptor
ALTERNATE_NAMES #formal.name Homo sapiens #common.name man
ORGANISM 10-Sep-1999 #sequence.revision 10-Sep-1999 #text.change
DATE 10-Sep-1999
ACCESSIONS A35356; A36475; A48416; A36007; A23666; B35010; I38094
REFERENCE A35356
#authors Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann,
#journal M.P.; Jertz, R.; Dower, S.K.; Cosman, D.; Goodwin, R.G.
#title Science (1990) 248:1019-1023
#cross-references MUID:90260639
#accession A35356
#status preliminary
#molecule_type mRNA
#residues 1-461 ##label SMI
#cross-references GB:M32315; NID:q189185; PIDN:AAA59929.1; PID:q189186
REFERENCE A36475
#authors Kohno, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King,
#journal M.W.; Hale, K.K.; Squires, C.H.; Thompson, R.C.; Vannice,
#title Proc. Natl. Acad. Sci. U.S.A. (1990) 87:8331-8335
#cross-references MUID:91045991
#accession A36475
#status preliminary
#molecule_type mRNA
#residues 1-195,'R',197-461 ##label KOH
#cross-references GB:M55994; GB:M38549; NID:g339757; PIDN:AAA6755.1;
#note PID:g339758
REFERENCE A48416
#authors Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.;
#journal Gentz, R.; Brockhaus, M.; Lesslauer, W.
#title Cytokine (1990) 2:231-237
#cross-references MUID:91370690
#accession A48416
#status preliminary
#molecule_type mRNA; protein
#residues 23-461 ##label DEM
#cross-references GB:S63368; NID:g235648; PIDN:AAI9824.1; PID:g235649
#note sequence extracted from NCBI backbone (NCBIN:63368,
#note NCBI:P:63371)
REFERENCE A36007
#authors Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang,
#journal D.; Ringold, G.M.
#title Proc. Natl. Acad. Sci. U.S.A. (1990) 87:6151-6155
#cross-references MUID:90349572
#accession A36007
#status preliminary
#molecule_type mRNA
#residues 116-140,'P',142-195,'R',197-362,'T',364-461 ##label HEL
#cross-references GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752
REFERENCE A23666
#authors Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.;
#journal Lesslauer, W.; Brockhaus, M.
#title J. Biol. Chem. (1990) 265:20131-20138
#cross-references MUID:91056048
#accession A23666
#status preliminary
#molecule_type protein
#residues 23-40;65-69;136-141;300-306 ##label LOE
REFERENCE A35010
#authors Engelmann, H.; Novick, D.; Wallach, D.

```

```

#journal J. Biol. Chem. (1990) 265:1531-1536
#title Two tumor necrosis factor-binding proteins purified from
#cross-references MUID:90110215
#accession B35010
#status preliminary
#molecule_type protein
#residues 27-31 ##label ENG
REFERENCE I38094
#authors Kunert, P.; Kemper, O.; Wallach, D.
#journal Gene (1994) 150:381-386
#title Cloning, sequencing and partial functional characterization
#cross-references MUID:95121934
#accession I38094
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-37 ##label RES
#cross-references EMBL:X80021; NID:g666044; PIDN:CAA56324.1;
#note PID:g825701
GENETICS
#gene GDB:TNFR2
#cross-references GDB:125914; OMIM:191191
#map_position 1p36.2-1p36.2
#introns 26/73
#note the list of introns is incomplete
CLASSIFICATION #superfamily tumor necrosis factor receptor type 2; NGF
#product repeat homology
FEATURE #duplication; glycoprotein; receptor; transmembrane protein
KEYWORDS #domain signal sequence #status predicted #label SIG\
#product tumor necrosis factor receptor type 2 #status
#experimental #label MAT\
40-76 #domain NGF receptor repeat homology #label NG1\
78-119 #domain NGF receptor repeat homology #label NG2\
120-162 #domain NGF receptor repeat homology #label NG3\
164-201 #domain NGF receptor repeat homology #label NG4\
262-279 #domain transmembrane #status predicted #label TMN\
280-461 #domain intracellular #status predicted #label INT\
171,193 #binding-site carbohydrate (Asn) (covalent) #status
#predicted
SUMMARY #length 461 #molecular-weight 48291 #checksum 5724
Query Match 5.5%; Score 95; DB 1; Length 461;
Best Local Similarity 42.4%; Pred. No. 7.45e-01;
Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;
Db 89 WPECLSCG-SRCSDDVETQACTREONRICRPG-WYCALSKOGRCRLC-APLRK-C 143
19 WLPESLSCNALCASP-V-SK-CLI-QE-LCQCRPEBGNCSCK-E-CMLCGLALDEC 70
RESULT 11
ENTRY #type complete
TITLE no1x protein - Rhizobium fredii
ORGANISM #formal.name Rhizobium fredii
DATE 03-Feb-1994 #sequence.revision 03-Feb-1994 #text.change
ACCESSIONS S35019
REFERENCE S35019
#authors Meinhardt, L.W.; Krishnan, H.B.; Baiatati, P.A.; Pueppke, S.G.
#journal Mol. Microbiol. (1993) 9:17-29
#title Molecular cloning and characterization of a sym plasmid locus
#cross-references MUID:94018604
#accession S35019
#molecule_type DNA
#residues 1-471 ##label MEI
#cross-references EMBL:L12251
#note the authors translated the codon GAA for residue 309 as

```

```

GENETICS      GLN
#gene
SUMMARY      #length 471 #molecular-weight 51205 #checksum 7158

Query Match      5.5%: Score 95; DB 2; Length 471;
Best Local Similarity 29.9%: Pred. No. 7,45e-01;
Matches 23; Conservative 21; Mismatches 30; Indels 3; Gaps 3;

Db 156 PSMPPDKAIEALDQ-PELFYAIGSGGRCGKITAKDLSEFKHNPQVAFQESQA 214
QY 83 SDPPPSKSTVEELHEIPSELRAL-TEGDTOLNMNIVSEPAEELSHENHIVSFLETVN 141
Db 215 QSYAQNT-IPSDSAENA 230
QY 142 QPHHQNVSPPSNVHAP 158

RESULT 12
ENTRY      MMRTS      #type complete
TITLE      laminin beta-2 chain precursor - rat
ALTERNATE_NAMES      laminin chain B3: S-laminin
ORGANISM      #formal_name Rattus norvegicus #common_name Norway rat
DATE      30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change
16-Jul-1999

ACCESSIONS      S03539
REFERENCE      Hunter, D.D.; Shah, V.; Merlie, J.P.; Sanes, J.R.
#authors      Nature (1989) 338:229-234
#journal      A laminin-like adhesive protein concentrated in the synaptic
#title      cleft of the neuromuscular junction.
#cross-references      M01D:89159410
#accession      S03539
#residues      1-1801 #label HUN
#molecule_type      mRNA
COMPLEX      ##cross-references      EMBL:X16563; NID:957250; PIDN:CA34561.1; PID:957251
#cross-references      EMBL:X16563; NID:957250; PIDN:CA34561.1; PID:957251
#residues      1-1801 #label HUN
#molecule_type      mRNA
FUNCTION      #description      Interact with cells and with other basement membrane proteins
to promote differentiation, development, and cell migration
CLASSIFICATION      #superfamily      laminin beta-1 chain; laminin-type EGF-like
homology
KEYWORDS      basement membrane; calcium binding; cell binding; coiled
coll.; extracellular matrix; glycoprotein; heptad repeat;
heterotrimer; nidogen binding
FEATURE      #domain      signal sequence #status predicted #label SIG
1-35
36-1801
#product      laminin beta-2 chain #status predicted #label
MAT
#domain      VI #label DOM6\
286-285 #domain V #label DOM5\
286-555 #domain laminin-type EGF-like homology #label LE01\
286-347 #domain laminin-type EGF-like homology #label LE02\
350-410 #domain laminin-type EGF-like homology #label LE03\
413-470 #domain laminin-type EGF-like homology #label LE04\
473-522 #domain laminin-type EGF-like homology #status atypical
525-555 #label LE05\
#domain      IV #label DOM4\
556-784 #domain laminin-type EGF-like homology #label LE06\
788-831 #domain IIT #label DOM3\
788-1196 #domain laminin-type EGF-like homology #label LE07\
834-877 #domain laminin-type EGF-like homology #label LE08\
880-927 #domain laminin-type EGF-like homology #label LE09\
930-966 #domain laminin-type EGF-like homology #label LE10\
989-1038 #domain laminin-type EGF-like homology #label LE11\
1041-1095 #domain laminin-type EGF-like homology #label LE12\
1098-1143 #domain laminin-type EGF-like homology #label LE13\
1146-1190 #domain IIT, heptad repeats #label DOM2\
1197-1412 #domain IIT, heptad repeats #label DOM2\
1413-1445 #domain alpha #label ALP\
1446-1801 #domain I, heptad repeats #label DOM1\
45-50 #disulfide_bonds #status predicted\
251,371,1088,1252,

```

```

1311,1351,1502 #binding_site carbohydrate (Asn) (covalent) #status
1193,1196,1800 #disulfide_bonds interchain #status predicted
SUMMARY      #length 1801 #molecular-weight 196472 #checksum 8852

```

```

Query Match      5.5%: Score 96; DB 1; Length 1801;
Best Local Similarity 36.7%: Pred. No. 5,45e-01;
Matches 11; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

```

```

Db 1041 CTCMLGTPQRCPTDLCRDPSTGQPC 1070
QY 26 CNKALCASDVSKCLIQELCQCRPGNCSC 55

```

```

RESULT 13
ENTRY      MMMSA      #type complete
TITLE      laminin alpha-1 chain precursor - mouse
ALTERNATE_NAMES      laminin chain A1
ORGANISM      #formal_name Mus musculus #common_name house mouse
DATE      30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change
16-Jul-1999

```

```

ACCESSIONS      A31771; A30449; S00624; A30450; S08895; S02678; S01790;
A30451; S14670
REFERENCE      A31771
#authors      Sasaki, M.; Kleinman, H.K.; Huber, H.; Deutzmann, R.; Yamada,
Y.
#journal      J. Biol. Chem. (1988) 263:16536-16544
#title      Laminin, a multidomain protein. The A chain has a unique
globular domain and homology with the basement membrane
proteoglycan and the laminin B chains.
#cross-references      M01D:89034134
#accession      A31771
#residues      1-3084 #label SAS
#molecule_type      mRNA

```

```

#cross-references      EMBL:J04064; NID:9309419; PIDN:AAA39410.1;
PID:9309420
#accession      A30449
#molecule_type      protein
#residues      183-195;570-571,'A',573-586;596-612,'X',614-617,'EMK';
630-646;1217-1222,'YPR',1226-1227;1303-1310;1498-1507;
2033-2040,'X',2042-2043;2137-2151;2156-2178;2227-2240;
2406-2420;2440-2451;2481-2486;2624-2639;2818-2843;
3009-3033,'V',3035 #label SA2

```

```

REFERENCE      S00624
#authors      Hartl, L.; Oberpaumer, J.; Deutzmann, R.
#journal      Eur. J. Biochem. (1988) 173:629-635
#title      The N terminus of laminin A chain is homologous to the B
chains.
#cross-references      M01D:88225080
#accession      S00624
#molecule_type      mRNA
#residues      1-208,'T',210-334 #label HAR
#cross-references      EMBL:X07737; NID:952857; PIDN:CAA30561.1; PID:952858
#accession      A30450
#molecule_type      protein
#residues      311-335,'N',337-339;630-642,'D',644;692-734;737-748,'X',
750-760,'G',762-763;765-769,'X',771,'R',773-778;
780-786,'X',788-802,'Q',839-852,'Q',854-855,'OXO',
859-869,'Q',871-874;1148-1158;1353-1389;1449-1459
#label HA2

```

```

#note      the sequence from Fig. 7 is inconsistent with that from
Fig. 5 in having 209-11e, 239-Thr, and 240-Arg; the
sequence from Fig. 7 is inconsistent with that from
Table 1 in having 335-Thr

```

```

REFERENCE      S08895
#authors      Mann, K.; Deutzmann, R.; Timpl, R.
#journal      Eur. J. Biochem. (1988) 178:71-80
#title      Characterization of proteolytic fragments of the
laminin-nidogen complex and their activity in
ligand-binding assays.
#cross-references      M01D:89078415
#accession      S08895
#molecule_type      protein

```



```
#authors          Saint-Jacques, E.; Seguin, C.  
#journal         DNA Cell Biol. (1993) 12:329-340  
#title           Cloning and nucleotide sequence of a complementary DNA  
                 encoding Xenopus laevis metallothionein: mRNA accumulation  
                 in response to heavy metals.  
#cross-references MIMD:93263990  
#accession      I51538  
##status        Preliminary; translated from GB/EMBL/DDBJ  
##molecule-type mRNA  
##residues     1-62 ##label SAM  
##cross-references GB:M6729; NID:g214585; PIDN:AAB59949.1; PID:g214586  
CLASSIFICATION #superfamily metallothionein  
SUMMARY        #length 62 #molecular-weight 6403 #checksum 3001
```

Query Match 5.4%; Score 94; DB 2; Length 62;  
Best Local Similarity 44.4%; Pred. No. 1.0e+00;  
Matches 12; Conservative 6; Mismatches 7; Indels 2; Gaps 2;

DB 38 CPACSKC-SQG-CHEKSGSKSCCN 62  
ON 31 CASDVSKLDELCCCRPGEGNGSCG 57  
::: ||| | :| :|||:  
|:: ||| | :| :|||:

RESULT 15

ENTRY S14458 #type complete  
TITLE laminin alpha-1 chain precursor - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change  
13-Aug-1999  
S14458: S14663; A34961  
S14458 Haaparanta, T.; Uitto, J.; Ruoslahti, E.; Engvall, E.  
#journal Matrix (1991) 11:1151-1150  
#title Molecular cloning of the cDNA encoding human laminin A chain  
#cross-references MIMD:9133420  
#accession S14458  
##status not compared with conceptual translation  
##molecule-type mRNA  
##residues 1-3075 ##label HAA

REFERENCE S14663  
#authors Nissinen, M.; Vuolteenaho, R.; Boot-Handford, R.; Kallunki,  
P.; Tryggvason, K.  
#journal Biochem. J. (1991) 276:369-379  
#title Primary structure of the human laminin A chain. Limited  
expression in human tissues.  
#cross-references MIMD:91264789  
#accession S14663  
##molecule-type mRNA  
##residues 1-227,'FE','230-251','MLP','255-418','E','420-518','L',  
520-1022,'V','1024-1074','V','1076-1339','W','1341-1512,  
'P','1514-2078','KV','2081-2628 ##label NIS  
'P',1514-2078,'KV','2081-2628' NID:934225; PIDN:CAGI418.1; PID:934226

REFERENCE A34961  
##cross-references EMBL:X58531; NID:934225; PIDN:CAGI418.1; PID:934226  
#authors Olesen, D.; Nagayoshi, T.; Fazio, M.; Peltonen, J.; Jaakkola,  
S.; Sanborn, D.; Sasaki, T.; Kuivaniemi, H.; Chu, M.L.;  
Deutermann, R.; Timpl, R.; Uitto, J.  
#journal Lab. Invest. (1989) 60:772-782  
#title Human laminin: cloning and sequence analysis of CDNA  
encoding A, B1 and B2 chains, and expression of the  
corresponding genes in human skin and cultured cells.  
#cross-references MIMD:89280632  
#accession A34961  
##status not compared with conceptual translation  
##molecule-type mRNA  
##residues 'W','2397-2745','L','2747-3053','L','3055-3072','PSP' ##label  
#note OLS  
The authors translated the codon AGA for residue 2692 as pro

GENETICS  
#gene GDB:IAMAI: LAMA  
##cross-references GDB:120135; OMIM:150320  
#map position 18p11.32-18p11.22  
CLASSIFICATION #superfamily laminin alpha-1 chain; laminin G repeat

```

KEYWORDS
    homology: laminin-type EGF-like homology
    basement membrane: calcium binding; cell binding; coiled
    coil; disulfide bond; extracellular matrix; glycoprotein;
    heparin binding; heptad repeat; heterotrimer

FEATURE
1-17
18-3075

    #domain signal sequence #status predicted #label SIG\
    #product laminin alpha-1 chain #status predicted #label
    MARY
    #domain VI #label DOM6\
    #domain V #label DOM5\
    #domain laminin-type EGF-like homology #label LE1\
    #domain laminin-type EGF-like homology #label LE2\
    #domain laminin-type EGF-like homology #label LE3\
    #domain laminin-type EGF-like homology #label LE4\
    #domain laminin-type EGF-like homology #status atypical\
    #label LE5\
    #domain IVB #label DO4B\
    #domain IIB #label DO3B\
    #domain laminin-type EGF-like homology #status atypical\
    #label LE6\
    #domain laminin-type EGF-like homology #label LE7\
    #domain laminin-type EGF-like homology #label LE8\
    #domain laminin-type EGF-like homology #label LE9\
    #domain laminin-type EGF-like homology #label LE10\
    #domain laminin-type EGF-like homology #label LE11\
    #domain laminin-type EGF-like homology #label LE12\
    #domain laminin-type EGF-like homology #label LE13\
    #domain laminin-type EGF-like homology #status atypical\
    #label LE14\
    #domain laminin-type EGF-like homology #status atypical\
    #label LE15\
    #domain laminin-type EGF-like homology #status atypical\
    #label LE16\
    #domain IIVa #label DO4A\
    #domain IIVa #label DO3A\
    #domain laminin-type EGF-like homology #status atypical\
    #label LE17\
    #domain laminin-type EGF-like homology #label LE18\
    #domain laminin-type EGF-like homology #label LE19\
    #domain laminin-type EGF-like homology #label LE20\
    #domain I/II, heptad repeats #label DOM2\
    #region cell attachment #status predicted\
    #domain G #label DOMG\
    #domain laminin G repeat homology #label LG1\
    #domain laminin G repeat homology #label LG2\
    #domain laminin G repeat homology #label LG3\
    #region cell attachment (R-G-D) motif\
    #domain laminin G repeat homology #label LG4\
    #domain laminin G repeat homology #label LG5\
    38,144,555,665,763,
    801,838,926,952,
    1045,1407,1579,
    1596,1678,1689,
    1698,1717,1804,
    1894,1898,1957,
    1974,1991,2038,
    2047,2094,2098,
    2243,2244,2384,
    2408,2518,2519,
    2729,2852,2915,
    2983

    #binding-site carbohydrate (Asn) (covalent) #status
    predicted\
    #disulfide bonds #status predicted
    #length 3075 #molecular-weight 337155 #checksum 7424

SUMMARY
    297-305
    Query Match 5.4%, Score 94, DB 2, Length 3075;
    Best Local Similarity 35.0%; Pred. No. 1,01e+00;
    Matches 14; Conservative 8; Mismatches 16; Indels 2; Gaps 2;

DB 1084 FPDGVPDODDGRSGDAGNEEGGLGCVETGACPC-KE 1122
      ||:||||:||||:||||:||||:||||:||||:||||:
QY 20 LPESLCKKALCADSVKRC-LIQELCCQPRGEGSCSCKE 58

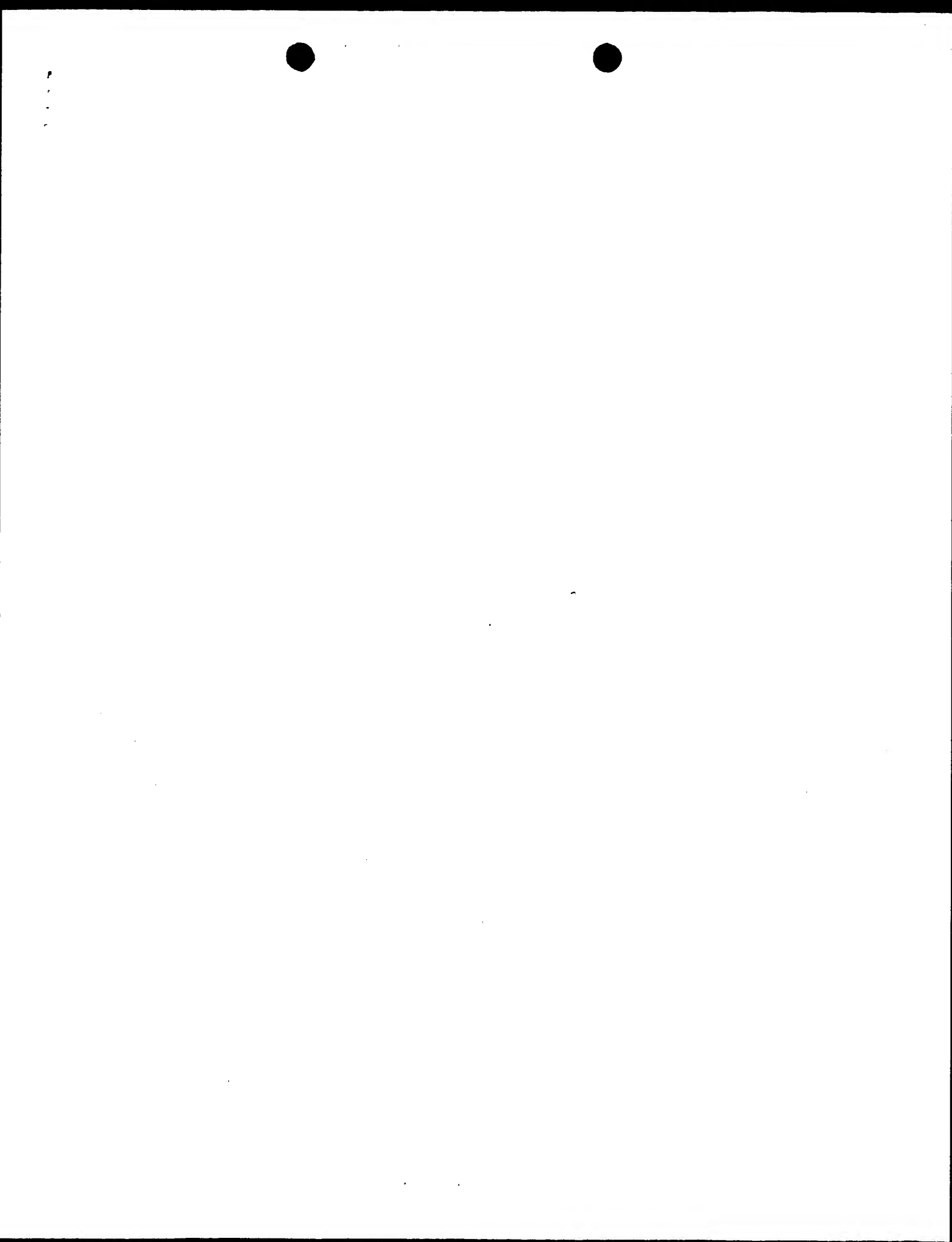
```

Tue Aug 29 09:10:18 2000

US-09-416-267-2.rpr

Page 9

Search completed: Wed Aug 16 09:31:22 2000  
Job time : 79 secs.





Db	1	MLKHYAVLTLALIMFLTMPELPSISCKNALCASVSKCLIQELCQCPBEGNCCCKCKM	60
Qy	1	MKLYHAVLTLALIMFLTMPELPSISCKNALCASVSKCLIQELCQCPBEGNCCCKCKM	60
Db	61	LCGLALMDECCDVGKCNCRNRSDFPRSKSTVEELHEPILPSLFRALTEBQDLWNINVS	120
Qy	61	LCGLALMDECCDVGKCNCRNRSDFPRSKSTVEELHEPILPSLFRALTEBQDLWNINVS	120
Db	121	FPVAEELSHENLVSEFLTVNQPHQNVSVPSNNVAHPVSSDKEHNCTVYFDDCMSTHQ	180
Qy	121	FPVAEELSHENLVSEFLTVNQPHQNVSVPSNNVAHPVSSDKEHNCTVYFDDCMSTHQ	180
Db	181	CKTSCESMGASKYRMFNHNAACEGTPGECIDYGSKTYKCMKCMF	223
Qy	181	CKTSCESMGASKYRMFNHNAACEGTPGECIDYGSKTYKCMKCMF	223



DE Sequence 2, Application PC/TUS9603935  
XX  
CC Sequence 2, Application PC/TUS9603935  
CC GENERAL INFORMATION:  
CC APPLICANT: LI, YI  
CC APPLICANT: SU, KUI  
CC TITLE OF INVENTION: HUMAN CYTOKINE POLYPEPTIDE  
CC NUMBER OF SEQUENCES: 9  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
CC ADDRESSEE: STEWART & OLSTEIN  
CC STREET: 6 BECKER FARM ROAD  
CC CITY: ROSELAND  
CC STATE: NEW JERSEY  
CC COUNTRY: USA  
CC ZIP: 07068-1739  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US96/03935  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ferraro, Gregory D  
CC REGISTRATION NUMBER: 36,134  
CC REFERENCE/DOCKET NUMBER: 325800-566  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 201-994-1700  
CC TELEFAX: 201-994-1744  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 223 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 223 AA; 25017 MW; 262453 CN;  
SQ  
Query Match 100.0%; Score 1738; DB 1; Length 223;  
Best Local Similarity 100.0%; Pred. No. 6,64e-185;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MKLHYAVVLTLLIMFLTWLPESLSCNKALCASDVSKLIQELCQCRPGEGNSCKCKECM 60  
1 MKLHYAVVLTLLIMFLTWLPESLSCNKALCASDVSKLIQELCQCRPGEGNSCKCKECM 60  
DB 61 LCLGALWDECCDVGKNCNPNRNTSDTPPTSKSTVEELHEPIPSLFRALTEGDTOLNNIVS 120  
QY 61 LCLGALWDECCDVGKNCNPNRNTSDTPPTSKSTVEELHEPIPSLFRALTEGDTOLNNIVS 120  
DB 121 FPVAEELSHENLVSELETYNOPHONVSPSNNVHAPYSSDKHEHMCVTYYVFDCKMSIQ 180  
QY 121 FPVAEELSHENLVSELETYNOPHONVSPSNNVHAPYSSDKHEHMCVTYYVFDCKMSIQ 180  
DB 181 CKICESMGASKYRWFHNAACCECIGPECIDYGSRTYKCNKCMF 223  
QY 181 CKICESMGASKYRWFHNAACCECIGPECIDYGSRTYKCNKCMF 223  
RESULT 5  
ID US-08-843-651-2 STANDARD; PRT; 223 AA.  
AC xxxxxx  
XX  
DT  
XX  
XX Sequence 2, Application US/08843651  
XX  
CC Sequence 2, Application US/08843651  
CC GENERAL INFORMATION:  
CC APPLICANT: Holtzman, Douglas

CC TITLE OF INVENTION: NOVEL PEPTIDES WITHIN THE  
CC TITLE OF INVENTION: GROWTH FACTOR SUPERFAMILY  
CC NUMBER OF SEQUENCES: 2  
CC CORRESPONDENCE ADDRESSES:  
CC ADDRESSEE: Fish & Richardson P.C.  
CC STREET: 225 Franklin Street  
CC CITY: Boston  
CC STATE: MA  
CC COUNTRY: US  
CC ZIP: 02110-2804  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Diskette  
CC COMPUTER: IBM Compatible  
CC OPERATING SYSTEM: Windows95  
CC SOFTWARE: FASTSEQ for Windows Version 2.0  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/843,651  
CC FILING DATE: 16-APR-1997  
CC PRIORITY APPLICATION DATA:  
CC APPLICATION NUMBER:  
CC FILING DATE:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Melkielejohn, Ph.D., Anita L.  
CC REGISTRATION NUMBER: 35,283  
CC REFERENCE/DOCKET NUMBER: 09404/027001  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617/542-5070  
CC TELEFAX: 617/542-8906  
CC TELEX: 200154  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 223 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC FRAGMENT TYPE: internal  
CC SEQUENCE 223 AA; 25017 MW; 262453 CN;  
SQ  
Query Match 100.0%; Score 1738; DB 14; Length 223;  
Best Local Similarity 100.0%; Pred. No. 6,64e-185;  
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MKLHYAVVLTLLIMFLTWLPESLSCNKALCASDVSKLIQELCQCRPGEGNSCKCKECM 60  
1 MKLHYAVVLTLLIMFLTWLPESLSCNKALCASDVSKLIQELCQCRPGEGNSCKCKECM 60  
DB 61 LCLGALWDECCDVGKNCNPNRNTSDTPPTSKSTVEELHEPIPSLFRALTEGDTOLNNIVS 120  
QY 61 LCLGALWDECCDVGKNCNPNRNTSDTPPTSKSTVEELHEPIPSLFRALTEGDTOLNNIVS 120  
DB 121 FPVAEELSHENLVSELETYNOPHONVSPSNNVHAPYSSDKHEHMCVTYYVFDCKMSIQ 180  
QY 121 FPVAEELSHENLVSELETYNOPHONVSPSNNVHAPYSSDKHEHMCVTYYVFDCKMSIQ 180  
DB 181 CKICESMGASKYRWFHNAACCECIGPECIDYGSRTYKCNKCMF 223  
QY 181 CKICESMGASKYRWFHNAACCECIGPECIDYGSRTYKCNKCMF 223  
RESULT 6  
ID US-09-015-412-7 STANDARD; PRT; 223 AA.  
AC xxxxxx  
XX  
DT  
XX  
XX Sequence 7, Application US/09015412  
XX  
CC Sequence 7, Application US/09015412  
CC GENERAL INFORMATION:  
CC APPLICANT: Hillman, Jennifer L.  
CC APPLICANT: Lal, Preeti  
CC APPLICANT: Corley, Neil C.

CC	Sequence 437, Application US/09330781
CC	GENERAL INFORMATION:
CC	APPLICANT: Gearing, David P.
CC	APPLICANT: Robison, Keith E.
CC	APPLICANT: Holzman, Douglas A.
CC	TITLE OF INVENTION: Nucleic Acid Molecules Derived from a
CC	TITLE OF INVENTION: Human Prostate Stromal Cell Library
CC	FILE REFERENCE: MLN98-269A
CC	CURRENT APPLICATION NUMBER: US/09/330,781
CC	CURRENT FILING DATE: 1999-06-11
CC	EARLIER APPLICATION NUMBER: 60/090,179
CC	EARLIER FILING DATE: 1998-06-22
CC	NUMBER OF SEQ ID NOS: 490
CC	SOFTWARE: FastSeq for Windows Version 3.0
CC	SEQ ID NO 437
CC	LENGTH: 113
CC	TYPE: PRT
CC	ORGANISM: Homo sapiens
CC	FEATURE:
CC	NAME/KEY: SIGNAL
CC	LOCATION: (1)..(25)
CC	NAME/KEY: VARIANT
CC	LOCATION: (1)..(113)
CC	OTHER INFORMATION: Xaa = Any Amino Acid
CC	SEQUENCE 113 AA; 12537 MW; 64990 CN;
CC	Query Match 48.2%; Score 838; DB 19; Length 113;
CC	Best Local Similarity 98.1%; Pred. No. 8,13e-80;
CC	Matches 105; Conservative 1; Mismatches 1; Indels 0; Gaps 0
CC	Db 1 MKLHYAVLTLALMLFELTPESLSCNKLALCASDVSKCLTIOELQCRPEGNCSCKRECM 60
CC	1 MKLHYAVLTLALMLFELTPESLSCNKLALCASDVSKCLTIOELQCRPEGNCSCKRECM 60
CC	Oy 61 LCLGALMDECCDCVGMCPNPNRSTDPXTSKSTVEELHDPISLFLAL 107
CC	61 LCLGALMDECCDCVGMCPNPNRSTDPXTSKSTVEELHDPISLFLAL 107
CC	Oy 61 LCLGALMDECCDCVGMCPNPNRSTDPXTSKSTVEELHDPISLFLAL 107
CC	RESULT 8
CC	ID US-60-169-629-463 STANDARD; PRT; 125 AA.
CC	AC xxxxxx
CC	XX
CC	XX
CC	XX
CC	DE
CC	XX
CC	Sequence 463, Application US/60169629
CC	GENERAL INFORMATION:
CC	APPLICANT: Dumas Milne Edwards, J.B.
CC	APPLICANT: Bouquellet, L.
CC	APPLICANT: Jobert, S.
CC	TITLE OF INVENTION: CDNAS for Secreted Proteins
CC	FILE REFERENCE: GENSET.071PRT
CC	CURRENT APPLICATION NUMBER: US/60/169,629
CC	CURRENT FILING DATE: 1999-12-08
CC	NUMBER OF SEQ ID NOS: 715
CC	SOFTWARE: Patent.pm
CC	SEQ ID NO 463
CC	LENGTH: 125
CC	TYPE: PRT
CC	ORGANISM: Homo sapiens
CC	FEATURE:
CC	NAME/KEY: SIGNAL
CC	LOCATION: -14..-1
CC	SEQUENCE 125 AA; 13870 MW; 87972 CN;
CC	Query Match 43.8%; Score 761; DB 3; Length 125;
CC	Best Local Similarity 93.8%; Pred. No. 5,91e-71;
CC	Matches 106; Conservative 2; Mismatches 4; Indels 1; Gaps 1

QY 60 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 119  
DB 61 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKGN-CQQLTF 112  
QY 120 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKHEMCTVYVF 172

RESULT 9  
ID US-60-187-470-463 STANDARD: PRT: 125 AA.

Sequence 463, Application US/60187470

GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Bouqueloret, L.

FILE REFERENCE: 78.052.PRO

CURRENT APPLICATION NUMBER: US/60/187,470

CURRENT FILING DATE: 2000-03-06

NUMBER OF SEQ ID NOS: 715

SOFTWARE: Patent.pm

SEQ ID NO 463

LENGTH: 125

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: SIGNAL

LOCATION: -14...-1

SEQUENCE 125 AA: 13870 MW; 87972 CN;

Query Match 43.8%; Score 761; DB 3; Length 125;  
Best Local Similarity 93.8%; Pred. No. 5,91e-71;

Matches 106; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

DB 1 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 60

QY 60 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 119

QY 61 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKGN-CQQLTF 112

QY 120 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKHEMCTVYVF 172

RESULT 10  
ID US-60-197-873-14313 STANDARD: PRT: 128 AA.

Sequence 14313, Application US/60197873

GENERAL INFORMATION:  
APPLICANT: Bejanin, Stephanie

APPLICANT: Tanaka, Hiroaki

APPLICANT: Dumas Milne Edwards, Jean Baptiste

APPLICANT: Jobert, Severin

FILE REFERENCE: 81.051.PRO

CURRENT APPLICATION NUMBER: US/60/197,873

NUMBER OF SEQ ID NOS: 52153

SOFTWARE: Patent.pm

CC LENGTH: 128  
CC TYPE: PRT  
CC ORGANISM: Homo sapiens  
CC FEATURE:  
CC NAME/KEY: SIGNAL  
CC LOCATION: -14...-1  
SQ SEQUENCE 128 AA: 14367 MW; 96202 CN;

Query Match 43.8%; Score 761; DB 23; Length 128;  
Best Local Similarity 93.8%; Pred. No. 5,91e-71;

Matches 106; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

DB 1 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 60

QY 60 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 119

DB 61 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKGN-CQQLTF 112

QY 120 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKHEMCTVYVF 172

RESULT 11  
ID US-60-147-499-4240 STANDARD: PRT: 128 AA.

Sequence 4240, Application US/60147499

GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/60/147,499

CURRENT FILING DATE: 1999-08-05

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 4240

LENGTH: 128

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: SIGNAL

LOCATION: -14...-1

SEQUENCE 128 AA: 14367 MW; 96202 CN;

Query Match 43.8%; Score 761; DB 3; Length 128;  
Best Local Similarity 93.8%; Pred. No. 5,91e-71;

Matches 106; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

DB 1 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 60

QY 60 MLCIGALWDECCDCVGMCPNRYSDTPTSKSTVEELHEPIPSLFRALTEGDTOLNMNIV 119

DB 61 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKGN-CQQLTF 112

QY 120 SFPVAEELSHHENLVSFLETYNOPHHONVSPSNNVHAPYSSDKHEMCTVYVF 172

RESULT 12  
ID US-09-307-140-1146 STANDARD: PRT: 75 AA.

Sequence 1146, Application US/09307140

Sequence 1146, Application US/09307140

CC GENERAL INFORMATION:  
CC APPLICANT: Gearling, David P.  
CC APPLICANT: Holtzman, Douglas A.  
CC APPLICANT: Robison, Keith E.  
CC TITLE OF INVENTION: Nucleic Acid Molecules Derived from a  
CC FILE REFERENCE: MN98-14pa  
CC CURRENT APPLICATION NUMBER: US/09/307,140  
CC CURRENT FILING DATE: 1999-05-07  
CC EARLIER APPLICATION NUMBER: 60/084,565  
CC EARLIER FILING DATE: 1998-05-07  
CC NUMBER OF SEQ ID NOS: 1168  
CC SOFTWARE: FastSeq for Windows Version 3.0  
CC SEQ ID NO 1146  
CC LENGTH: 75  
CC TYPE: PRT  
CC ORGANISM: Homo sapiens  
CC FEATURE:  
CC NAME/KEY: SIGNAL  
CC LOCATION: (1)...(25)  
SQ SEQUENCE 75 AA; 8284 MW; 24423 CN;  
Query Match 35.2%; Score 612; DB 19; Length 75;  
Best Local Similarity 100.0%; Pred. No. 6,31e-54;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 MKLHYAVLTALIMFLTWLPESLSCNKLACASDVSKCLIOELCCCRGEGNCGCKRCM 60  
QY 1 MKLHYAVLTALIMFLTWLPESLSCNKLACASDVSKCLIOELCCCRGEGNCGCKRCM 60  
Db 61 LCLGALMDECCDCV 74  
QY 61 LCLGALMDECCDCV 74  
RESULT 13  
ID US-08-820-970-9 STANDARD; PRT; 206 AA.  
AC xxxxxx  
DT  
DT  
DT  
DT  
DE Sequence 9, Application US/08820970  
CC GENERAL INFORMATION:  
CC APPLICANT: LI, YI  
CC APPLICANT: SU, KUI  
CC APPLICANT: LI, HAODONG  
CC TITLE OF INVENTION: HUMAN CYTOKINE POLYPEPTIDE  
CC NUMBER OF SEQUENCES: 9  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
CC STREET: 9410 KEY WEST AVENUE  
CC CITY: ROCKVILLE  
CC STATE: MARYLAND  
CC COUNTRY: USA  
CC ZIP: 20850  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: Patentin Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/820,970  
CC FILING DATE: 20-MAR-1997  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Brookes, Anders, A.  
CC REGISTRATION NUMBER: 36,373  
CC REFERENCE/DOCKET NUMBER: PF270  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 301-309-8504

CC TELEFAX: 301-309-8512  
CC INFORMATION FOR SEQ ID NO: 9:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 206 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 206 AA; 22216 MW; 187648 CN;  
Query Match 31.2%; Score 542; DB 14; Length 206;  
Best Local Similarity 35.4%; Pred. No. 5,34e-46;  
Matches 79; Conservative 55; Mismatches 70; Indels 19; Gaps 14;  
Db 1 MOLICYFVILFVGIAFW-SSLANDGCGNMGVSGVSKCLITQSCCKRLND--CHCKKDC 57  
QY 1 MKLH-YVAVLTALIMFLTWLPESLSCNKLACASDVSKCLIOELCCCRGEGNCGCKRC 59  
Db 58 LNLGELTYECGGLDMC-PKHQDVLPSTLPSEIGDI-EGVPELFDTLADDE-GNST 114  
QY 60 MLCGLALMDECCDCVGCNFRNYSIDTP-TSKSYELHEPIPSLFRALTEGDTLNNMI 118  
Db 115 IRESMRAGEKOR--VA--QG-GAS--GDAGNGNGNGAG-SAGYT-LCTVYI-NSCIRA 163  
QY 119 VSFPAVEELSHHNLVSLFETVQPHQVSVPSNNVHAPYSSDEHMKTYVTFDDCSI 178  
Db 164 NKRCQCESMGASSTYWFHDGCGCEVGENCLYGINESRCRC 206  
QY 179 HCKICSGESMGASKYRMFHNACECIGPICIDYGSXTVCNMC 221  
RESULT 14  
ID PCT-US96-03935-9 STANDARD; PRT; 206 AA.  
AC xxxxxx  
DT  
DT  
DT  
DT  
DE Sequence 9, Application PC/TUS9603935  
CC GENERAL INFORMATION:  
CC APPLICANT: LI, YI  
CC APPLICANT: SU, KUI  
CC TITLE OF INVENTION: HUMAN CYTOKINE POLYPEPTIDE  
CC NUMBER OF SEQUENCES: 9  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,  
CC STREET: STEWART & OLSTEIN  
CC CITY: ROSELAND  
CC STATE: NEW JERSEY  
CC COUNTRY: USA  
CC ZIP: 07068-1739  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC OPERATING SYSTEM: IBM PC compatible  
CC SOFTWARE: Patentin Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US96/03935  
CC FILING DATE:  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Ferraro, Gregory D  
CC REGISTRATION NUMBER: 36,134  
CC REFERENCE/DOCKET NUMBER: 325800-566  
CC TELECOMMUNICATION INFORMATION:  
CC TELEFAX: 201-994-1744  
CC TELEPHONE: 201-994-1700  
CC INFORMATION FOR SEQ ID NO: 9:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 206 amino acids

CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 206 AA; 22216 MW; 187648 CN;

Query Match 31.2%; Score 542; DB 1; Length 206;  
Best Local Similarity 35.4%; Pred. No. 5.34e-46;  
Matches 79; Conservative 55; Mismatches 70; Indels 19; Gaps 14;

Db 1 MQLLCFYILFVGIAPW-SSLANDGCNEVYCGSVSKCLITOSCCKLND--CHCKKDC 57  
QY 1 MKLH-YVAVLTLATLIMFLTWLPESLSCNKALCASDVSKCLIQELCCCRPGEGNCSCCKEC 59  
Db 58 LNCGLGELYIECCGCLDMC-PKHKDVLPSTLRPSRIGDI-EGVPELFDTLTAEDDE-GWST 114  
QY 60 MLCIGALMBECCDCVGCNPRNYSPTPP-TSKSTVEELHEPIPSLFRALTEGDTQLNMNI 118  
Db 115 IFRSMRAGFKR--VA--QG-GAS--GDAGNGNGNAG-SAGVT-LCTVIY-NSCIRA 163  
QY 119 VSFVPAEELSHENILVSFLQVNPQHONVSVPSNNVHAPYSSDKRHMCTVYVFDQMSI 178  
Db 164 NKCRQCESMGASSYRWFHDGCECEVGENCLNNGINESRCRC 206  
QY 179 HQCKISCSMGASKYRWFHNAACECIGPECIDYGSKTVCAMC 221

RESULT 15  
ID US-09-416-267-9 STANDARD; PRT: 206 AA.  
XX XXXXXX  
XX DT  
XX

Sequence 9, Application US/09416267

CC Sequence 9, Application US/09416267  
CC GENERAL INFORMATION:  
CC APPLICANT: LI, YI  
CC APPLICANT: SU, KUI  
CC APPLICANT: LI, HAODONG  
CC TITLE OF INVENTION: HUMAN CYTOKINE POLYPEPTIDE  
CC NUMBER OF SEQUENCES: 9  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Human Genome Sciences, Inc.  
CC STREET: 9410 Key West Avenue  
CC CITY: Rockville  
CC STATE: MD  
CC COUNTRY: USA  
CC ZIP: 20850  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/09/416,267  
CC FILING DATE: OCT-12-1999  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Joseph J. Kenny  
CC REGISTRATION NUMBER: 43,710  
CC REFERENCE/DOCKET NUMBER: PF270P1  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 301-309-8504  
CC TELEFAX: 301-309-8439  
CC INFORMATION FOR SEQ ID NO: 9:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 206 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein

SQ SEQUENCE 206 AA; 22216 MW; 187648 CN;  
Query Match 31.2%; Score 542; DB 20; Length 206;  
Best Local Similarity 35.4%; Pred. No. 5.34e-46;  
Matches 79; Conservative 55; Mismatches 70; Indels 19; Gaps 14;

Db 1 MQLLCFYILFVGIAPW-SSLANDGCNEVYCGSVSKCLITOSCCKLND--CHCKKDC 57  
QY 1 MKLH-YVAVLTLATLIMFLTWLPESLSCNKALCASDVSKCLIQELCCCRPGEGNCSCCKEC 59  
Db 58 LNCGLGELYIECCGCLDMC-PKHKDVLPSTLRPSRIGDI-EGVPELFDTLTAEDDE-GWST 114  
QY 60 MLCIGALMBECCDCVGCNPRNYSPTPP-TSKSTVEELHEPIPSLFRALTEGDTQLNMNI 118  
Db 115 IFRSMRAGFKR--VA--QG-GAS--GDAGNGNGNAG-SAGVT-LCTVIY-NSCIRA 163  
QY 119 VSFVPAEELSHENILVSFLQVNPQHONVSVPSNNVHAPYSSDKRHMCTVYVFDQMSI 178  
Db 164 NKCRQCESMGASSYRWFHDGCECEVGENCLNNGINESRCRC 206  
QY 179 HQCKISCSMGASKYRWFHNAACECIGPECIDYGSKTVCAMC 221

Search completed: Wed Aug 16 09:40:22 2000  
Job time : 277 secs.





Db 1 MKLHYAVLTALITLMTLPESLSCNKALCASDVSKLLOELCCRRPBGNCSCCKEEM 60  
 QY 1 MKLHYAVLTALITLMTLPESLSCNKALCASDVSKLLOELCCRRPBGNCSCCKEEM 60  
 Db 61 LCLGALMDECCDCVGMCPNRYSDPTPTSKSTVEELHEPIPSLFRALTEGDTQLNMNIVS 120  
 QY 61 LCLGALMDECCDCVGMCPNRYSDPTPTSKSTVEELHEPIPSLFRALTEGDTQLNMNIVS 120  
 Db 121 FPVAEELSHHENIVSELTYNQPHONVSPSNVHAPYSDDKREKCTVYVFDQMSIHQ 180  
 QY 121 FPVAEELSHHENIVSELTYNQPHONVSPSNVHAPYSDDKREKCTVYVFDQMSIHQ 180  
 Db 181 CRISCEMGAASKYRMFHNACCECIGPECIDYGSKTVKCMNCF 223  
 QY 181 CRISCEMGAASKYRMFHNACCECIGPECIDYGSKTVKCMNCF 223

RESULT 2  
 ID W0591 standard; Protein; 223 AA.  
 AC W0591;  
 DT 15-FEB-1999 (first entry)  
 Human growth factor Tango-67.  
 Tango-67; growth factor; human; inflammation; cell proliferation;  
 therapy; diagnosis; wound healing; tissue repair; vlnetary.  
 OS Homo sapiens.  
 PN W09846641-A1.  
 PD 22-OCT-1998.  
 PR 15-APR-1998; U07603.  
 PR 16-APR-1997; US-843651.  
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
 PI Holzman D;  
 DR WPI; 98-609915/51.  
 DR N-PSDB; V63465.  
 PT New isolated growth factor, Tango-67 - used to develop products for  
 the diagnosis and treatment of inflammation or disorders associated  
 PT with cellular proliferation  
 PS Claim 5; Fig 1; 69pp; English.  
 CC This is the amino acid sequence of Tango-67, a novel human soluble  
 CC growth factor which is capable of promoting cell proliferation  
 CC and/or differentiation. The sequence was deduced from the  
 CC nucleotide sequence (see V63465) of a cDNA clone obtained from a  
 CC human astrocyte cDNA library. Tango-67 is related to a number of  
 CC growth factors, particularly members of the connective tissue  
 CC growth factor family such as TSG. The invention provides Tango-67  
 CC nucleic acid molecules and polypeptides, host cells, vectors and  
 CC antibodies. The products can be used for the treatment of  
 CC disorders associated with aberrant expression or activity of  
 CC Tango-67. In particular they can be used to diagnose or treat  
 CC inflammation or disorders associated with cellular proliferation.  
 CC The Tango-67 polypeptides can be used in wound healing, tissue  
 CC repair, implant fixation, or stimulation of bone growth. The  
 CC products can also be used for detection, diagnosis and screening  
 CC assays.  
 CC Sequence 223 AA:  
 SQ

Query Match 100.0%; Score 1738; DB 1; Length 223;  
 Best Local Similarity 100.0%; Pred. No. 8.73e-158;  
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 CRISCEMGAASKYRMFHNACCECIGPECIDYGSKTVKCMNCF 223  
 ID Y13362 standard; Protein; 420 AA.  
 AC Y13362;  
 DT 25-JUN-1999 (first entry)  
 Amino acid sequence of protein PRO214.  
 DE Secreted protein; transmembrane protein; human; enterocolitis;  
 KW Zollinger-Ellison syndrome; gastrointestinal ulceration;  
 KW congenital microvillus atrophy; skin disease; cell growth;  
 KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;  
 KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;  
 KW fibromodulin; dermal scarring; Usher Syndrome; Atrophla areata;  
 KW anti-thrombotic; wound healing; tissue repair.  
 OS Homo sapiens.  
 PN W09914328-A2.  
 PD 25-MAR-1999.  
 PR 16-SEP-1998; U19330.  
 PR 25-NOV-1997; US-066840.  
 PR 17-SEP-1997; US-059113.  
 PR 17-SEP-1997; US-059115.  
 PR 17-SEP-1997; US-059117.  
 PR 17-SEP-1997; US-059119.  
 PR 17-SEP-1997; US-059121.  
 PR 17-SEP-1997; US-059122.  
 PR 17-SEP-1997; US-059184.  
 PR 18-SEP-1997; US-059263.  
 PR 18-SEP-1997; US-059266.  
 PR 15-OCT-1997; US-062125.  
 PR 17-OCT-1997; US-062285.  
 PR 17-OCT-1997; US-062287.  
 PR 21-OCT-1997; US-063486.  
 PR 24-OCT-1997; US-062814.  
 PR 24-OCT-1997; US-062816.  
 PR 24-OCT-1997; US-063045.  
 PR 24-OCT-1997; US-063120.  
 PR 24-OCT-1997; US-063121.  
 PR 24-OCT-1997; US-063127.  
 PR 24-OCT-1997; US-063128.  
 PR 27-OCT-1997; US-063329.  
 PR 27-OCT-1997; US-063327.  
 PR 28-OCT-1997; US-063541.  
 PR 28-OCT-1997; US-063542.  
 PR 28-OCT-1997; US-063544.  
 PR 28-OCT-1997; US-063549.  
 PR 28-OCT-1997; US-063550.  
 PR 28-OCT-1997; US-063564.  
 PR 29-OCT-1997; US-063435.  
 PR 29-OCT-1997; US-063704.  
 PR 29-OCT-1997; US-063732.  
 PR 29-OCT-1997; US-063738.  
 PR 29-OCT-1997; US-063734.  
 PR 29-OCT-1997; US-064215.  
 PR 29-OCT-1997; US-064215.  
 PR 31-OCT-1997; US-063870.  
 PR 31-OCT-1997; US-064109.  
 PR 03-NOV-1997; US-064248.  
 PR 07-NOV-1997; US-064809.  
 PR 12-NOV-1997; US-065186.  
 PR 17-NOV-1997; US-065186.  
 PR 18-NOV-1997; US-065693.  
 PR 21-NOV-1997; US-066120.  
 PR 21-NOV-1997; US-066364.  
 PR 24-NOV-1997; US-066772.  
 PR 24-NOV-1997; US-066466.  
 PR 24-NOV-1997; US-066770.  
 PR 24-NOV-1997; US-066511.  
 PR 24-NOV-1997; US-066453.  
 PA (GETH) GENENTECH INC.  
 PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;  
 DR WPI; 99-229533/19.  
 DR N-PSDB; X52233.



PT New isolated human genes and polypeptides used in, e.g. treatment of gastrointestinal ulceration.

PS Claim 12: Fig 40; 320pp; English.

CC Y13344-403 represent secreted and transmembrane human proteins.

CC The cDNA sequences are obtained from cDNA libraries, prepared from fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.

CC The encoded polypeptides have specific uses based on their homology to known polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated with the preservation and maintenance of gastrointestinal mucosa and the repair of acute and chronic mucosal lesions (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as lung squamous cell carcinoma of the vulva and gliomas), potent effects on cell growth and development, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophia areata; PRO269 can be used as an anti-thrombotic agent; PRO287 polypeptides and portions may have therapeutic applications in wound healing and tissue repair; PRO317 can be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, e.g. in the heart of genital tract.

CC Sequence 420 AA;

DB 236 WALTHIKVDIDEGTEGANGADFCVNTGSEYECRCARACIGCMGA 284

QY 19 WPEISLSC-NKALCASDVSKLIQELCCRPGEKNC-SCCKEMLCIGA 65

Query Match 5.8%; Score 100; DB 1; Length 420; Best Local Similarity 28.6%; Pred. No. 8.28e+00; Matches 14; Conservative 12; Mismatches 21; Indels 2; Gaps 2;

RESULT 4

ID Y05281 standard; Protein; 420 AA.

AC Y05281;

DT 22-JUN-1999 (first entry)

DE EGF-like homologue PRO214.

KW EBAF-2; inhibitor; tumour growth; cancer; EGF-like homologue;

KM EGF-8 homologue.

OS Homo sapiens.

PI WO9914337-A2.

PR 25-MAR-1999;

PR 10-SEP-1998; U18824.

PR 25-NOV-1997; US-066840.

PR 17-SEP-1997; US-059114.

PR 18-SEP-1997; US-059117.

PR 15-OCT-1997; US-062125.

PR 17-OCT-1997; US-062285.

PR 17-OCT-1997; US-062287.

PR 24-OCT-1997; US-062816.

PR 29-OCT-1997; US-063704.

PA (GETH) GENENTECH INC.

PI Botstein D, Goddard A, Gurney A, Hillan K, Lawrence DA, Roy M, Wood WJ.

DR WPI; 99-229532/19.

DR N-PSDB; X28431.

PT Antibodies against specific proteins overexpressed in tumours

PS Example 1; Fig 10; 130pp; English.

CC This sequence represents the EGF-like homologue PRO214.

CC The invention relates to antibodies (Ab) that bind to any of the polypeptides (1) designated PRO187; PRO533; PRO214; PRO240; PRO211; PRO230; PRO261 or EBAF-2. The Ab, or other agents that inhibit expression and/or activity of (1) are used: (1) to inhibit growth of tumours; and (11) as diagnostic/prognostic reagents for detection or quantification of (1) in cells or tissues, by standard immunoassays, with overexpression being indicative of cancer. For therapeutic use, the Ab may be conjugated to a toxin, chemotherapeutic agent or radioisotope.

CC Genes expressing (1), many of which are growth factor homologues, are

CC overexpressed in some cases of cancer.

SQ Sequence 420 AA;

Query Match 5.8%; Score 100; DB 1; Length 420; Best Local Similarity 28.6%; Pred. No. 8.28e+00; Matches 14; Conservative 12; Mismatches 21; Indels 2; Gaps 2;

DB 236 WALTHIKVDIDEGTEGANGADFCVNTGSEYECRCARACIGCMGA 284

QY 19 WPEISLSC-NKALCASDVSKLIQELCCRPGEKNC-SCCKEMLCIGA 65

Query Match 5.6%; Score 97; DB 1; Length 3084; Best Local Similarity 40.0%; Pred. No. 1.38e+01; Matches 16; Conservative 5; Mismatches 17; Indels 2; Gaps 2;

DB 1091 FPDCLPCPCHLPTLPHTCDLGGICSCSEDSGTGSC-KE 1129

QY 20 IPESLSCNKALCASDVSKC-LIQELCCRPGEKNCSCCKE 58

RESULT 6

ID W94641 standard; peptide; 77 AA.

AC W94641;

DT 29-APR-1999 (first entry)

DE TNF-R extracellular Cys-rich domain TNF-R p75

KW Tumour necrosis factor receptor; TNF-R; autoimmune diseases; inflammation; septic shock; cachexia; graft versus host disease; skin allergic reaction; immune complex disease; malaria;

KM Homo sapiens.

OS Homo sapiens.

PI WO9853842-A1.

PN 03-DEC-1998.

PR 29-MAY-1998; U10891.

PR 30-MAY-1997; US-86545.

PA (UYPE-) UNIV PENNSYLVANIA.

PI Greene MI, Murail R, Takasaki W;

DR WPI; 99-080781/07.

PT New compounds designed from a binding loop of a tumour necrosis factor receptor - are capable of inhibiting the biological activities of tumour necrosis factor, e.g., in treating inflammation or autoimmune diseases

PT Disclosure; Fig 1; 78pp; English.

CC The present invention describes peptides and peptide analogues which correspond in primary sequence to a binding loop of a tumour necrosis

CC factor receptor (TNF-R) superfamily member. The compounds are especially  
 CC designed from a binding loop of TNF-R p55. They are capable of  
 CC inhibiting TNF binding to its cellular receptors and may be used to  
 CC inhibit the biological activities of TNF. They may be used in treating  
 CC TNF-associated conditions such as acute and chronic inflammatory  
 CC responses, septic shock, cachexia, autoimmunity, graft-versus-host  
 CC disease, skin allergic reactions, immune complex disease, e.g. oral,  
 CC transplantation rejection and malaria. Administration is, e.g. oral,  
 CC intramuscular, intravenous, subcutaneous, intravenous or  
 CC intramuscular. Parenteral dosage is 0.1-5 mg/kg/day. The present  
 CC sequence represents an extracellular Cys-rich domain of TNF-R from the  
 CC present invention.  
 SQ Sequence 77 AA.

Query Match 5.5%; Score 95; DB 1; Length 77;  
 Best Local Similarity 42.4%; Pred. No. 1.93e+01;  
 Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;

DB 16 WPECLSCG-SRCSDDYETQACTREONRICRPG-WYCALSKOEGRLC-APLRK-C 70  
 19 WPECLSKKALCASP-V-SK-CLT-QE-LCCPRGEGNCSCCK-E-CMLCLGALMDEC 70

RESULT 7  
 ID W52271 standard; Protein; 122 AA.

AC W52271:  
 DE 29-JUN-1998 (first entry)  
 KW Soluble tumour necrosis factor receptor; sTNFR; TNF-mediated disease;  
 KW tumour necrosis factor binding protein; autoimmunity disease; arthritis;  
 KW adult respiratory distress syndrome; cachexia/anorexia; cancer; therapy;  
 KW chronic fatigue syndrome; graft rejection; Alzheimer's disease; TNBP.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 1..31  
 FT /note="optionally deleted"  
 FT Misc\_difference 32..115  
 FT /note="residues 32-115 of human TNF inhibitor"  
 FT Misc\_difference 116..122  
 FT /note="optionally deleted"  
 FT W09801555-A2.  
 PN 15-JAN-1998.  
 PD 09-JUL-1997; U12244.  
 PR 04-MAR-1997; US-039792.  
 PR 09-JUL-1996; US-021443.  
 PR 06-DEC-1996; US-032534.  
 PR 23-JAN-1997; US-037737.  
 PR 07-FEB-1997; US-039314.  
 PA (AMGE-) AMGEN INC.  
 PI Edwards CK, Fisher EF, Kieft GL;  
 WPI: 98-101052/09.

PT Truncated and soluble forms of tumour necrosis factor receptor -  
 PT useful for treating diseases involving factor, e.g. arthritis and  
 PT adult respiratory distress syndrome  
 PS Claim 3; Page -: 205pp; English.  
 CC This sequence is a truncated soluble tumour necrosis factor receptor  
 CC (sTNFR) protein of the invention. The truncated sTNFR proteins and tumour  
 CC necrosis factor binding proteins (TNBP) are used to treat any  
 CC TNF-mediated disease, e.g. arthritis, adult respiratory distress  
 CC syndrome, cachexia/anorexia, cancer, chronic fatigue syndrome, graft  
 CC rejection, Alzheimer's disease and other autoimmune diseases. Cells  
 CC transformed with a vector containing DNA encoding the protein may be used  
 CC for production of recombinant sTNFR, which may also be used for measuring  
 CC the amount of sTNFR in samples and to raise antibodies against sTNFR.  
 CC TNBP may also be used in preparation of therapeutic compositions for  
 CC treating the above diseases. The sTNFR proteins are well suited to large  
 CC scale production (since they lack the deamidation site in region 111-126,  
 CC so are more stable in vivo); contain fewer disulphide bonds and fewer  
 CC epitopes, making them less antigenic than full-length proteins.  
 SQ Sequence 122 AA;

Query Match 5.5%; Score 95; DB 1; Length 122;  
 Best Local Similarity 42.4%; Pred. No. 1.93e+01;

Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;  
 DB 67 WPECLSCG-SRCSDDYETQACTREONRICRPG-WYCALSKOEGRLC-APLRK-C 121  
 19 WPECLSKKALCASP-V-SK-CLT-QE-LCCPRGEGNCSCCK-E-CMLCLGALMDEC 70

RESULT 8  
 ID R77421 standard; Protein; 183 AA.

AC R77421:  
 DE 10-FEB-1996 (first entry)  
 DE BamP delta53 nerve growth factor sequence.  
 KW Nerve growth factor; neurotrophic factor; therapeutic;  
 KW protein refolding; NGF; plasmid pT3X-2.  
 OS Synthetic.  
 PN W09530686-A1.  
 PD 16-NOV-1995.  
 PR 02-MAY-1995; U05423.  
 PR 09-MAY-1994; US-240122.  
 PR 27-JUN-1994; US-265080.  
 PA (SYNT) SYNTEX-STRONGEN NEUROSCIENCE JOINT VENTU.  
 PI Bonam D, Kohno T, Lile J, Rosendahl MS;  
 WPI: 95-404080/51.  
 DR N-PSDB; T05443.  
 PT Process for bacterial expression of recombinant neurotrophic factor  
 PT - useful for promoting the survival and maintaining phenotypic  
 PT differentiation of nerve and glial cells.  
 PS Example 1; Page 36-37; 57pp; English.  
 CC The synthetic nerve growth factor (NGF) gene isolated from Bam TP  
 CC delta 53 plasmid pT3X1-2 is designed to optimize codons for  
 CC expression in Escherichia coli as well as create unique sites for  
 CC subsequent cloning steps. The recombinant protein is solubilized  
 CC and sulfonlated and allowed to refold in the presence of PEG and  
 CC urea. Biologically active NGF, used for promoting the survival of  
 CC and maintaining the phenotypic differentiation of nerve and glial  
 CC cells, is isolated and purified. This method breaks incorrectly  
 CC formed disulphide bonds and allows refolding of the factor into  
 CC the correct tertiary structure required for maximum yield of full  
 CC active protein.  
 SQ Sequence 183 AA;

Query Match 5.5%; Score 95; DB 1; Length 183;  
 Best Local Similarity 42.4%; Pred. No. 1.93e+01;  
 Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;

DB 68 WPECLSCG-SRCSDDYETQACTREONRICRPG-WYCALSKOEGRLC-APLRK-C 122  
 19 WPECLSKKALCASP-V-SK-CLT-QE-LCCPRGEGNCSCCK-E-CMLCLGALMDEC 70

RESULT 9  
 ID W52270 standard; Protein; 235 AA.

AC W52270:  
 DE 29-JUN-1998 (first entry)  
 KW Soluble tumour necrosis factor receptor; sTNFR; TNF-mediated disease;  
 KW tumour necrosis factor binding protein; autoimmunity disease; arthritis;  
 KW adult respiratory distress syndrome; cachexia/anorexia; cancer; therapy;  
 KW tumour necrosis factor inhibitor; Alzheimer's disease; TNBP.  
 OS Homo sapiens.  
 PN W09801555-A2.  
 PD 15-JAN-1998.  
 PD 09-JUL-1997; U12244.  
 PR 04-MAR-1997; US-039792.  
 PR 09-JUL-1996; US-021443.  
 PR 06-DEC-1996; US-032534.  
 PR 23-JAN-1997; US-037737.  
 PR 07-FEB-1997; US-039314.  
 PA (AMGE-) AMGEN INC.  
 PI Edwards CK, Fisher EF, Kieft GL;  
 WPI: 98-101052/09.  
 DR N-PSDB; V19802.  
 PT Truncated and soluble forms of tumour necrosis factor receptor -

PT useful for treating diseases involving factor, e.g. arthritis and adult respiratory distress syndrome

PS Claim 3; Fig 8; 205pp; English.

CC This sequence is the human tumour necrosis factor inhibitor. The protein was used to make the truncated soluble tumour necrosis factor receptor (sTNFR) proteins of the invention. The truncated sTNFR proteins and tumour necrosis factor binding proteins (TNBP) are used to treat any TNF-mediated disease, e.g. arthritis, adult respiratory distress syndrome, cachexia/anorexia, cancer, chronic fatigue syndrome, graft rejection, Alzheimer's disease and other autoimmune diseases. Cells transformed with a vector containing DNA encoding the protein may be used for production of recombinant sTNFR, which may also be used for measuring the amount of sTNFR in samples and to raise antibodies against sTNFR. TNBP may also be used in preparation of therapeutic compositions for treating the above diseases. The sTNFR proteins are well suited to large scale production (since they lack the deamidation site in region 111-126, so are more stable in vivo) contain fewer disulphide bonds and fewer epitopes, making them less antigenic than full-length proteins.

Sequence 235 AA.

Query Match 5.5%; Score 95; DB 1; Length 235;  
Best Local Similarity 42.4%; Pred. No. 1.93e+01;  
Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;

DB 67 WPECLSCG-SRCSDDVETACTREONRICRPG-WYCALSKOGGRLC-APLRK-C 121  
QY 19 WLPESLSCNKALCASP-V-SK-CLI-QE-LCCPRGEGNCSCCK-E-CMLCIGALMDEC 70

RESULT 10  
ID W59665 standard; Protein; 235 AA.  
AC W59665;

DE 28-SEP-1998 (first entry)  
DE Human soluble tumour necrosis factor receptor type II.  
KW Human; tumour necrosis factor; TNF; TNF receptor type II;  
KW inflammatory disease; leukaemia; TNF binding protein;  
KW anti-inflammatory drug; methotrexates.

OS Homo sapiens.  
PN W09824463-A2.

PD 11-JUN-1998.  
PR 08-DEC-1997; U22733.  
PR 09-JUL-1997; US-052023.  
PR 06-DEC-1996; US-032587.  
PR 23-JAN-1997; US-036355.  
PR 07-FEB-1997; US-039315.

(AMGE-) AMGEN INC.  
Bendele AM, Edwards CK, Sennello RM;  
DR N-PSDB; V41549.

PT Treatment of acute or chronic inflammatory disease, e.g. leukaemia - by administering tumour necrosis factor binding protein and at least one additional anti-inflammatory drug, e.g. methotrexate

PS Disclosure; Fig 2; 104pp; English.

CC This is the amino acid sequence of the human tumour necrosis factor receptor type II, used in the method of the invention involving the treatment of acute or chronic inflammatory disease such as leukaemia by administering tumour necrosis factor binding protein and at least one additional anti-inflammatory drug, e.g. methotrexate.

Sequence 235 AA;

Query Match 5.5%; Score 95; DB 1; Length 235;  
Best Local Similarity 42.4%; Pred. No. 1.93e+01;  
Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;

DB 67 WPECLSCG-SRCSDDVETACTREONRICRPG-WYCALSKOGGRLC-APLRK-C 121  
QY 19 WLPESLSCNKALCASP-V-SK-CLI-QE-LCCPRGEGNCSCCK-E-CMLCIGALMDEC 70

RESULT 11  
ID W89234 standard; Protein; 235 AA.  
AC W89234;  
DT 04-MAR-1999 (first entry)

DE Tumour necrosis inhibitor 40 kDa protein.  
KW Tumour necrosis factor receptor 1; TNFR-1; inhibitor; osteoprotegerin;  
KW OPG; chimeric; fusion; dimerisation domain; autoimmune disease;  
KW inflammation; apoptosis.

OS Homo sapiens.  
PN W09849305-A1.

PD 05-NOV-1998.  
PR 29-APR-1998; U08631.  
PR 01-MAY-1997; US-850188.

PA (AMGE-) AMGEN INC.  
Boyle WJ, Wooden S;  
DR WPI; 99-034661/03.

PT New chimeric osteoprotegerin polypeptides - contain the osteoprotegerin dimerisation domain and a heterologous sequence, useful to treat TNF and TNFR-mediated disorders

PS Disclosure; Fig 3; 92pp; English.

CC The present invention describes a chimeric polypeptide (A1), comprising an osteoprotegerin (OPG) dimerisation domain fused to a heterologous amino acid sequence. Also described are: (1) a multimer polypeptide comprising covalently associated A1 monomers; (2) an isolated nucleic acid encoding A1; (3) an expression vector comprising the nucleic acid sequence; and (4) a host cell transformed or transfected with the expression vector so that the nucleic acid is expressible. The products from the present invention are useful to treat a variety of disorders including those related to receptor binding. Compositions comprising tumour necrosis factor (TNF)/OPG and TNF receptor (TNFR)/OPG chimeras are used to treat TNF and TNFR-mediated disorders such as inflammation, autoimmune diseases and disorders related to excessive apoptosis. The chimeras are also useful for detecting molecules which interact with fused heterologous sequences to identify potential new receptors and CC ligands. The present sequence represents the TNF inhibitor 40 kDa protein.

Sequence 235 AA;

Query Match 5.5%; Score 95; DB 1; Length 235;  
Best Local Similarity 42.4%; Pred. No. 1.93e+01;  
Matches 25; Conservative 10; Mismatches 13; Indels 11; Gaps 11;

DB 67 WPECLSCG-SRCSDDVETACTREONRICRPG-WYCALSKOGGRLC-APLRK-C 121  
QY 19 WLPESLSCNKALCASP-V-SK-CLI-QE-LCCPRGEGNCSCCK-E-CMLCIGALMDEC 70

RESULT 12  
ID R72504 standard; Protein; 461 AA.  
AC R72504;

DE 31-OCT-1995 (first entry)  
DE p75 Tumour Necrosis Factor Receptor.  
KW Ligand; tumour necrosis factor; nerve growth factor; TNF; NGF;  
KW receptor.

OS Homo sapiens.

FH Key Location/Qualifiers  
FT region 27..214  
FT domain /label= TBPII.  
FT 258..285

FT /label= Transmembrane domain.  
FT msc\_difference 259 /note= "Unidentified amino acid."

PN EP-648783-A.

PD 19-APR-1995.  
PR 11-OCT-1994; 116015.  
PR 12-OCT-1993; IL-107267.

PA (YEDA ) YEDA RES & DEV CO LTD.  
PA (WALL/) WALLACH D.  
PI Beletsky I, Bigda J, Mett I, Wallach D;

DR N-PSDB; Q89544.

PT Tumour necrosis factor (TNF) receptor ligand - used to increase inhibitory effect of a soluble TNF receptor

PS Disclosure; Figure 2; 18pp; English.

CC A ligand to a member of the tumour necrosis factor (TNF)/nerve growth factor (NGF) receptor family which binds either to the region





